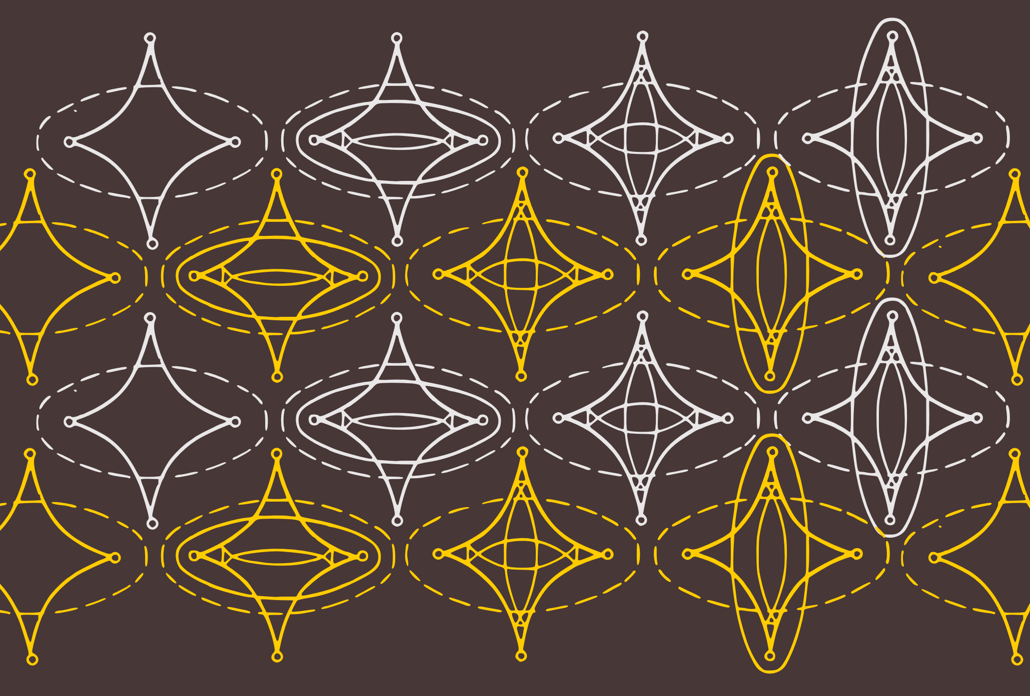
V. Arnold

Chapitres Supplémentaires de la Théorie des Équations Différentielles Ordinaires



Editions Mir Moscu

V. ARNOLD

CHAPITRES SUPPLÉMENTAIRES DE LA THÉORIE DES ÉQUATIONS DIFFÉRENTIELLES ORDINAIRES

ÉDITIONS MIR · MOSCOU

Traduit du russe par
DJILALI EMBAREK

На французском языке

TABLE DES MATIÈRES

AVANT-PROPOS

La principale découverte de Newton, celle qu'il a cru indispensable de garder en secret et qu'il n'a publiée que sous forme d'un anagramme, tient dans la phrase suivante: « *Data aequatione quotcunque fluentes quantitae involvente fluxiones invenire et vice versa.* » Ce qu'on traduirait dans le langage mathématique moderne par: « Il est utile de résoudre des équations différentielles ».

La théorie des équations différentielles représente actuellement un amalgame d'idées et de méthodes au plus haut point utiles aux applications les plus variées et moteur perpétuel des recherches théoriques dans toutes les branches des mathématiques.

Une grande partie des voies reliant les théories mathématiques abstraites avec les applications relève des équations différentielles. Certains domaines de la théorie des équations différentielles ont atteint un tel degré de développement qu'ils se sont érigés en sciences indépendantes; les problèmes de la théorie des équations différentielles ont grandement contribué à l'élaboration, par exemple, de l'algèbre linéaire, de la théorie des groupes de Lie, de l'analyse fonctionnelle, de la mécanique quantique, etc. Ainsi donc, les équations différentielles sont le pivot de la conception scientifique du monde.

L'auteur s'est attaché à construire son ouvrage autour des principes et méthodes fondamentaux utilisés dans l'étude des équations différentielles. Il a pris soin de ne pas alourdir de détails techniques les notions fondamentales qui par essence sont simples et suggestives. Les problèmes les plus importants et les plus simples sont traités avec minutie, tandis que les problèmes plus complexes sont sommairement exposés.

L'ouvrage commence par l'étude de certaines équations différentielles spéciales intégrables par quadrature. L'accent est mis non pas sur l'aspect formel de la théorie élémentaire d'intégration, mais sur ses liens avec les idées, méthodes et notions générales de mathématiques (résolution des singularités, groupes de Lie, dia-

grammes de Newton) d'une part, et les applications scientifiques, de l'autre.

La théorie des équations aux dérivées partielles du premier ordre est traitée à l'aide de la structure de contact naturelle dans la variété des 1-jets des fonctions. On expose au passage les éléments indispensables de la géométrie des structures de contact, qui rendent cette théorie indépendante d'autres sources.

Les méthodes dites qualitatives occupent une place importante dans cet ouvrage. Le développement actuel de la théorie qualitative des équations différentielles, élaborée par H. Poincaré, a conduit à l'idée que de même que l'intégration implicite des équations différentielles est généralement impossible, de même est impossible l'étude qualitative d'équations différentielles tant soit peu générales à espace des phases multidimensionnel. Cet ouvrage traite de l'analyse des équations différentielles en termes de stabilité structurelle, c'est-à-dire de stabilité des portraits de phases par rapport à de petites déformations des équations différentielles. Il passe en revue les principaux résultats obtenus dans ce domaine après les premiers travaux de A. Andronov et L. Pontriaguine: Fondements de la théorie des C-systèmes structurellement stables dont toutes les trajectoires sont exponentiellement instables, théorème de Smale de non-densité de l'ensemble des systèmes structurellement stables. Il traite également de l'impact de ces découvertes sur les applications (nous avons en vue la description de régimes chaotiques stables de mouvement du type turbulent).

Les méthodes asymptotiques comptent parmi les méthodes les plus puissantes et les plus souvent utilisées d'étude des équations différentielles. Cet ouvrage expose les idées principales de la méthode de moyennisation qui remonte aux travaux des fondateurs de la mécanique céleste et qui trouve de vastes applications dans tous les domaines nécessitant une séparation d'une évolution lente d'oscillations rapides (N. Bogolioubov, Y. Mitropolski et autres).

Malgré l'abondance des travaux sur la moyennisation, le problème de l'évolution présente des ombres même pour les plus simples des systèmes à plusieurs fréquences. Cet ouvrage passe en revue les travaux portant sur le passage par des résonances et sur la capture d'une trajectoire par une surface résonnante.

L'idée de la méthode de moyennisation consiste à éliminer les perturbations par un choix convenable du système de coordonnées. Cette même idée est à la base de la théorie des formes normales de Poincaré. La méthode des formes normales est la principale méthode de la théorie locale des équations différentielles, décrivant le comportement des trajectoires au voisinage d'un point singulier ou d'une trajectoire fermée. Cet ouvrage expose les éléments de la méthode des formes normales de Poincaré, y compris la démonstration du

théorème fondamental de Siegel de linéarisation d'une application holomorphe.

La méthode des formes normales de Poincaré est largement utilisée aussi bien dans l'étude d'une seule équation différentielle qu'en théorie des bifurcations qui considère toute une famille d'équations dépendant de paramètres.

La théorie des bifurcations étudie le comportement des trajectoires d'un système sous l'effet d'une variation des paramètres de ce dernier. Pour les valeurs génériques des paramètres on a généralement affaire à des systèmes génériques (tous les points singuliers sont simples, etc.). Mais si le système dépend de paramètres il présentera nécessairement des dégénérescences pour certaines valeurs des paramètres (par exemple, deux points singuliers du champ de vecteurs coïncideront).

Une famille générique à un paramètre ne donne lieu qu'à des dégénérescences simples (qu'il est impossible d'éliminer par une petite déformation de la famille). Il s'établit donc une hiérarchie des dégénérescences en fonction des codimensions des surfaces correspondantes dans l'espace fonctionnel de tous les systèmes étudiés : dans les familles génériques à un paramètre on ne rencontre que des dégénérescences correspondant à des surfaces de codimension un, etc.

L'application des idées et méthodes de la théorie générale des singularités des applications différentiables de Whitney a fait faire ces dernières années un important progrès à la théorie des bifurcations.

Ce livre s'achève par un chapitre sur la théorie des bifurcations. Dans ce chapitre on utilise les méthodes développées dans les chapitres précédents et l'on décrit les résultats obtenus dans ce domaine depuis les travaux fondamentaux de H. Poincaré et A. Andronov.

Dans son exposé l'auteur s'est attaché à éviter le style axiomatique et déductif dont le trait caractéristique est la présence de définitions non motivées masquant les idées et méthodes fondamentales; telles des paraboles, elles ne sont ensuite expliquées qu'aux disciples initiés.

L'axiomatisation et l'algébrisation de la mathématique qui, prétend-on, se poursuivent depuis déjà plus de 50 ans ont conduit à rendre illisibles un si grand nombre de textes mathématiques que la menace (qui a toujours pesé) de voir les mathématiques perdre contact avec la physique et les sciences naturelles est devenue réelle. L'auteur a agencé son exposé de façon à mettre ce livre à la portée non seulement des mathématiciens mais aussi de tous les usagers de la théorie des équations différentielles.

On suppose au lecteur les rudiments de mathématiques générales : espaces vectoriels, calcul différentiel et intégral, éléments de théorie des équations différentielles ordinaires. Il pourra consulter avec

profit le manuel de V. Arnold, Equations différentielles ordinaires, Editions Mir, 1974 *).

La conception de l'exposé permet au lecteur de sauter les passages qui lui ont semblé difficiles sans grand préjudice pour la suite: l'auteur a tenté dans la mesure du possible d'éviter les renvois d'un chapitre à l'autre et même d'un paragraphe à l'autre.

La base de ce livre est une série de cours facultatifs faits par l'auteur à la faculté de mécanique et de mathématiques de l'Université Lomonossov en 1970-1976 pour les étudiants des 2-ième et 3-ième années.

L'auteur remercie tous ceux qui ont contribué à l'amélioration du texte originel et tout particulièrement Y. Iliachenko, dont les notes ont été très utiles à la préparation de cet ouvrage, D. Anossov et V. Pliss, dont il a utilisé les remarques judicieuses.

V. Arnold

---

*) Pour l'exposé de certaines questions on se sert ou on rappelle les notions les plus élémentaires sur les formes différentielles, les groupes de Lie et les fonctions d'une variable complexe. Ces notions ne sont pas toutefois indispensables à la compréhension de la majeure partie du livre.

NOTATIONS UTILISÉES

$\mathbb{R}$	— ensemble des réels.
$\mathbb{C}$	— ensemble des complexes.
$\mathbb{Z}$	— ensemble des entiers.
$\mathbb{R}^n$	— espace vectoriel réel à n dimensions.
$\exists$	— quantificateur existentiel (il existe).
$\forall$	— quantificateur universel (quel que soit).
$a \in A$	— l'élément a appartient à l'ensemble A.
$A \subset B$	— A sous-ensemble de B
$A \cap B$	— intersection des ensembles A et B.
$A \cup B$	— réunion des ensembles A et B.
$A \setminus B$	— différence des ensembles A et B (partie de A non comprise dans B).
$A \times B$	— produit cartésien des ensembles A et B (ensemble des couples (a, b), $a \in A$, $b \in B$).
$A \oplus B$	— somme directe des espaces vectoriels A et B.
$f: A \to B$	— application f de A dans B.
$x \mapsto y$ ou $y = f(x)$	— l'application f associe à l'élément x l'élément y.
$\operatorname{Im} f$ ou $f(A)$	— image de l'application f (mais $\operatorname{Im} z$ et la partie imaginaire de z).
$f^{-1}(y)$	— antécédent du point y par l'application f (l'ensemble de tous les x tels que $f(x) = y$).
$\operatorname{Ker} f$	— noyau de l'opérateur linéaire f (antécédent complet de 0).
$\dot{f}$	— vitesse de variation de la fonction f (dérivée par rapport au temps t).
f', f_*, df/dx, Df/Dx	— dérivée de l'application f.
T_xM	— espace tangent à la variété M en x.
$A \Rightarrow B$	— A entraîne B.
$A \Leftrightarrow B$	— A équivalent à B.
$\omega_1 \wedge \omega_2$	— produit extérieur des formes différentielles ω_1 et ω_2.
$f \circ g$	— composition des applications $((f \circ g)(x) = f(g(x)))$.
◀, ▶	— début et fin d'une démonstration.
$L_v f$	— dérivée de la fonction f suivant la direction du champ de vecteurs v.

Soient $(x_1, \ldots, x_n)$ des fonctions coordonnées. Le vecteur v est alors défini par ses coordonnées $v_1, \ldots, v_n$. La dérivée suivant la direction du champ v est donnée par

$$L_v f = v_1 \partial f/\partial x_1 + \ldots + v_n \partial f/\partial x_n.$$

Dans un système de coordonnées fixes $(x_1, \ldots, x_n)$ on se sert des notations suivantes:

dx_k — fonction de vecteur, égale à la k-ième composante de ce vecteur.

$\frac{\partial}{\partial x_k}$ — champ de vecteurs dont la k-ième composante est égale à 1 et les autres à zéro.

Soit l'équation différentielle $\dot{x} = v(x)$. Le domaine de définition du second membre est appelé *espace des phases*, le point x, *point représentatif*, le vecteur $v(x)$, *vecteur vitesse*, v, *champ de vecteurs de la vitesse*. Si $x = \varphi(t)$ est solution de l'équation, on appelle *trajectoire* l'image de l'application φ, et *courbe intégrale* le graphe de l'application φ.

Pour l'équation différentielle $\dot{x} = v(x, t)$ le domaine de définition du second membre est dit *espace des phases élargi*; v définit un *champ de directions* dans l'espace des phases élargi; si $x = \varphi(t)$ est solution, le graphe de l'application φ est appelé *courbe intégrale*.

CHAPITRE PREMIER

EQUATIONS SPÉCIALES

Les méthodes de tous les domaines de la mathématique interviennent dans l'étude des équations différentielles. Ce chapitre traite de certaines équations et types d'équations spéciales. On souligne d'une part l'importance de ces équations pour les applications, et de l'autre, le lien des méthodes de recherche avec divers problèmes de mathématiques générales (éclatement des singularités, diagrammes de Newton, groupes de Lie de symétries, etc.). Ce chapitre s'achève par l'étude de la théorie élémentaire de l'équation stationnaire à une dimension de Schrödinger et la théorie géométrique de l'équation non linéaire du second ordre.

§ 1. Equations différentielles invariantes par des groupes de symétries

Dans ce paragraphe on se penchera sur les généralités servant de base aux méthodes d'intégration d'équations différentielles sous forme explicite. La théorie de la similitude, i.e. la théorie des équations homogènes et quasi homogènes, est traitée à titre d'exemple.

A. Groupes de symétries des équations différentielles.

Soit un champ de vecteurs v dans un espace des phases U.

Définition. On appelle *symétrie du champ* v un difféomorphisme $g\colon U \to U$ envoyant v dans lui-même:

$$v(gx) = g_{*x}v(x).$$

On dit alors que le champ v est invariant par le difféomorphisme g.

Exemple 1. Un champ de vecteurs défini sur un plan rapporté à un système de coordonnées (x, y) et dont les composantes ne dépendent pas de x est invariant par les translations le long de l'axe x (fig. 1).

Exemple 2. Un champ de vecteurs $x\partial_x + y\partial_y$ sur un plan euclidien (x, y) est invariant par les homothéties $g(x, y) = (\lambda x, \lambda y)$, et par les rotations.

Les symétries d'un champ forment un groupe.

Exercice. Trouver le groupe de symétries du champ $x\partial_x + y\partial$ sur un plan rapporté à un système de coordonnées (x, y).

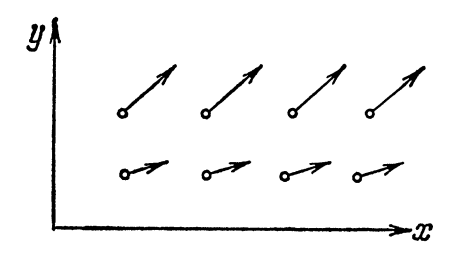


Fig. 1

Soit un champ de directions dans un espace des phases élargi.

Définition. On appelle *symétrie d'un champ de directions* un difféomorphisme de l'espace des phases élargi associant ce champ à lui-même. Autrement dit, le champ de directions est invariant par ce difféomorphisme.

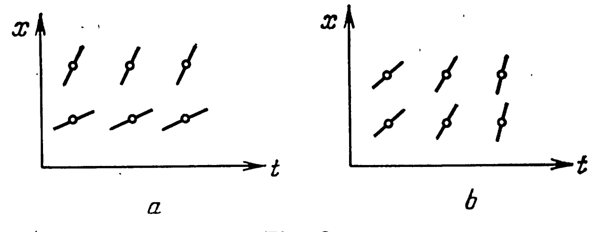


Fig. 2

Exemple 1. Le champ de directions de l'équation $\dot{x} = v(x)$ est invariant par toute translation le long de l'axe t (fig. 2 *a*).

Exemple 2. Le champ de directions de l'équation $\dot{x} = v(t)$ est invariant par toute translation le long de l'axe des x (fig. 2 *b*).

Définition. L'équation différentielle $\dot{x} = v(x)$ (resp. $\dot{x} = v(x, t)$) est *invariante* par un difféomorphisme g de l'espace des phases (resp. de l'espace des phases élargi) si le champ de vecteurs v (resp. le champ de directions v) l'est par le difféomorphisme g; on dit alors que g est une *symétrie* de cette équation.

Théorème. *Les trajectoires (resp. les courbes intégrales) d'une équation sont invariantes par toute symétrie de cette équation.* ▼

◀ Soient $x = \varphi(t)$ solution de l'équation $\dot{x} = v(x)$ et g, symétrie. Alors $x = g(\varphi(t))$ est également solution, puisque la symétrie envoie trajectoires dans trajectoires. La démonstration est analogue pour les courbes intégrales. ▶

Exemple. La famille des courbes intégrales de l'équation $\dot{x} = v(t)$ reste invariante par toute translation le long de l'axe des x, et celle de l'équation $\dot{x} = v(x)$, par toute translation le long de l'axe des t.

Les exemples suivants se rencontrent fréquemment dans les applications sous la dénomination de « théorie de similitude », « théorie des dimensions » ou encore « notions d'automodélisme ».

B. Equations homogènes.

Définition. Un champ de directions sur un plan privé du point O est *homogène* s'il est invariant par les homothéties

$$g^\lambda(x, y) = (e^\lambda x, e^\lambda y), \quad \lambda \in \mathbb{R},$$

L'équation différentielle $\frac{dy}{dx} = v(x, y)$ est *homogène* si son champ de directions l'est (fig. 3).

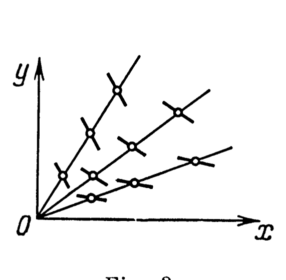


Fig. 3

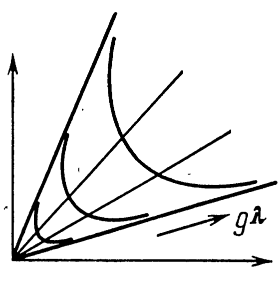


Fig. 4

Autrement dit, le champ doit avoir la même direction en tous les points des rayons issus de l'origine des coordonnées:

$$v(e^\lambda x, e^\lambda y) \equiv v(x, y).$$

Exemple. Une fonction f est homogène de degré d si $f(e^\lambda x, e^\lambda y) \equiv e^{\lambda d} f(x, y)$. Exemple: toute forme (tout polynôme homogène) de degré d. Soient P et Q deux formes de degré d de x et y. L'équation différentielle

$$\dot{x} = P, \quad \dot{y} = Q$$

est définie par un champ de vecteurs sur le plan. Le champ de directions correspondant dans le domaine $P \neq 0$ est champ de directions de l'équation homogène

$$\frac{dy}{dx} = \frac{Q}{P} \left(\text{par exemple, } \frac{dy}{dx} = \frac{ax + by}{cx + dy}, \ \frac{dy}{dx} = \frac{x^2 - y^2}{x^2 + y^2}, \text{ etc.} \right)$$

Remarque. Rien ne nous oblige à définir un champ homogène sur le plan tout entier privé du point O; on peut aussi bien envisager des champs homogènes sur un domaine homogène (= invariant par toute homothétie), par exemple, dans un angle de sommet O, etc.

T h é o r è m e. *Les courbes intégrales d'une équation homogène se déduisent l'une de l'autre par une homothétie* g^λ.

Donc, pour étudier une équation homogène il suffit de considérer une courbe intégrale dans chaque région du plan.

Ceci se démontre immédiatement grâce au théorème du point A.

E x e r c i c e. Prouver que les trajectoires du système $\dot{x} = P$, $\dot{y} = Q$, où P et Q sont des formes de degré d, se déduisent l'une de l'autre par homothétie (fig. 4).

Si l'une de ces trajectoires est fermée et parcourue en un temps T, l'homothétie g^λ lui associe une trajectoire fermée de période de révolution $\dfrac{T}{e^{\lambda(d-1)}}$.

C. Equations quasi homogènes et « calcul des dimensions ».

Fixons des réels α et β et considérons la famille de transformations:

$$g^s(x, y) = (e^{\alpha s}x, e^{\beta s}y). \qquad (1)$$

On remarquera que (1) définit un groupe à un paramètre de transformations linéaires du plan (fig. 5).

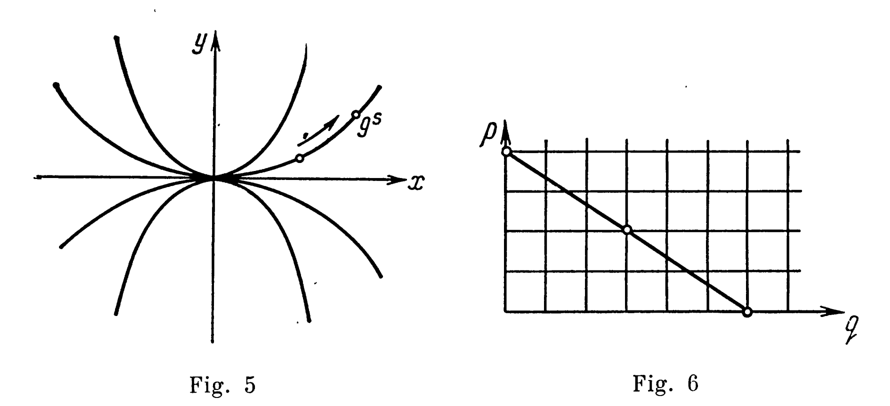


Fig. 5 Fig. 6

D é f i n i t i o n. On appelle *quasi homogène de degré* d une fonction f telle que

$$f(g^s(x, y)) \equiv e^{ds}f(x, y).$$

E x e m p l e. Si $\alpha = \beta = 1$, on obtient des fonctions homogènes ordinaires de degré d.

Les degrés de quasi homogénéité s'ajoutent dans la multiplication de fonctions. On les appelle également *poids*. Ainsi, x est de poids α, y de poids β, x^2y de poids $2\alpha + \beta$, etc. Les monômes quasi homogènes de degré fixe sont bien mis en évidence sur le *diagramme de Newton* de la figure 6. On représentera le monôme x^py^q par le

point (p, q) du réseau entier. Les puissances de tous les monômes de degré d sont les points entiers du segment d'équation $d = \alpha p + \beta q$ dans le plan (p, q).

Exercice. Pour quelles valeurs des poids la fonction $x^2 + xy^3$ est-elle quasi homogène?

Définition. L'équation différentielle $\frac{dy}{dx} = v(x, y)$ est *quasi homogène* (*de poids* α *et* β) si le champ de directions v est invariant par les transformations (1).

Le théorème général du point A entraîne le

Théorème. *Les courbes intégrales d'une équation quasi homogène se déduisent l'une de l'autre par les transformations* (1).

Exercice. Montrer que la fonction $v(x, y)$ définit une équation différentielle quasi homogène (de poids (α, β)) si, et seulement si, elle est quasi homogène de degré $d = \beta - \alpha$.

Remarque. Les définitions et théorèmes énoncés se généralisent sans peine à un nombre de variables supérieur à deux et aux équations différentielles d'ordre supérieur à un. En particulier, il est aisé de prouver le

Théorème. *Soient donnés dans le plan* (x, y) *une courbe* $\gamma: y = y(x)$ *et un point* (x_0, y_0) *en lequel* $\frac{d^ky}{dx^k} = F$. *Alors au point correspondant de la courbe* $g^s\gamma$ *on aura*

$$\frac{d^ky}{dx^k} = e^{(\beta - k\alpha)s}F.$$

En d'autres termes, d^ky/dx^k *se transforme par* (1) *comme* y/x^k, d'où l'intérêt de la notation d^ky/dx^k.

Exercice. Montrer que si dans le champ d'une force homogène de degré d une particule décrit une orbite Γ dans un temps T, elle décrira l'orbite homothétique $\lambda\Gamma$ dans le temps

$$T' = \lambda^{(1-d)/2}T.$$

Solution. L'équation de Newton $\frac{d^2x}{dt^2} = F(x)$, où $F(x)$ est homogène de degré d, est invariante par des applications convenables (1). Plus exactement, il faut prendre les poids α (pour x) et β (pour t) tels que $\alpha - 2\beta = \alpha d$, soit $\beta = \frac{1-d}{2}\alpha$. Donc l'orbite homothétique $x' = \lambda x$ sera parcourue pendant $T' = \lambda^{(1-d)/2}T$.

Exercice. Démontrer la troisième loi de Kepler: les carrés des temps de parcours d'orbites homothétiques sont entre eux comme les cubes des dimensions linéaires.

Solution. En faisant $d = -2$ (loi de la gravitation universelle) dans la solution de l'exercice précédent, on obtient $T' = \lambda^{3/2}T$.

Exercice. Etablir une relation entre l'amplitude de la période d'oscillations dans le cas d'une force de rappel proportionnelle à l'écart (oscillateur linéaire) et le cube de l'écart (force molle).

R é p o n s e. La période d'un pendule linéaire ne dépend pas de l'amplitude, celle d'un pendule mou est inversement proportionnelle à l'amplitude.

E x e r c i c e. On sait qu'une toupie à axe vertical possède une vitesse angulaire critique: si la vitesse angulaire est supérieure à la vitesse critique, la toupie occupe une position verticale stable, si elle est inférieure, la toupie tombe.

Comment varie la vitesse angulaire critique si l'on transporte la toupie sur la Lune où l'accélération de la pesanteur est six fois moindre que sur Terre?

R é p o n s e: La vitesse diminue de $\sqrt{6}$ fois.

D. Applications des groupes à un paramètre de symétries à l'abaissement de l'ordre.

T h é o r è m e. *Si l'on connaît un groupe à un paramètre de symétries d'un champ de directions dans $\mathbb{R}^n$, l'intégration de l'équation*

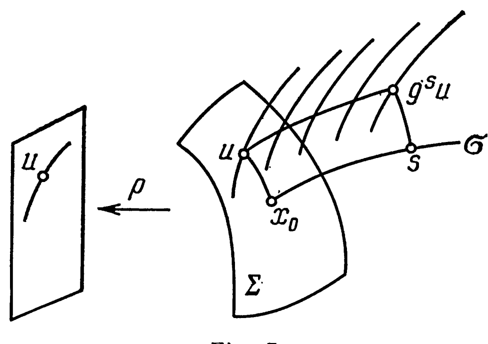


Fig. 7

différentielle correspondante dans $\mathbb{R}^n$ se ramène à une intégration dans $\mathbb{R}^{n-1}$.

En particulier, si l'on connaît un groupe à un paramètre de symétries d'un champ de directions sur le plan, l'équation correspondante $\frac{dy}{dx} = f(x, y)$ s'intègre explicitement.

◀ Soit $\{g^s\}$ le groupe de symétries envisagé. Considérons les orbites $\{g^s x\}$ du flot $\{g^s\}$. On peut (au moins, localement) définir un *espace d'orbites* $n - 1$ dimensionnel (l'espace quotient par l'action de g^s) et une application p de l'espace initial sur l'espace quotient (les orbites du flot $\{g^s\}$ sont envoyées dans des points par l'application p). Il se trouve que le champ de directions initial se projette par l'application p dans un champ de directions dans l'espace d'orbites $n - 1$ dimensionnel; il ne reste plus qu'à intégrer le champ obtenu. ▶

Plus exactement, considérons un point quelconque $x_0 \in \mathbb{R}^n$ (fig. 7); soit σ l'orbite du groupe de symétries $\{g^s\}$ passant par x_0. Menons par le point x_0 transversale locale Σ de $n - 1$ dimensions à la courbe σ. Traçons au voisinage du point x_0 un système de coordonnées locales (s, u), où au couple $s \in \mathbb{R}$, $u \in \Sigma$ est associé le point $g^s u$ de l'espace initial. La projection p sur l'espace

des orbites et l'action du groupe de symétries g^s sont données au voisinage du point x_0 par

$$p(s, u) = u, \quad g^{s_1}(s_2, u) = (s_1 + s_2, u)$$

(les points de la surface Σ paramétrisent les orbites locales).

A noter que si l'on connaît explicitement le groupe g^s, on peut calculer explicitement les coordonnées (s, u). Ecrivons l'équation donnée dans ces coordonnées. Si le champ de directions n'est pas tangent à la surface Σ au point x_0 (ce que l'on peut toujours réaliser par le choix de Σ), alors au voisinage de ce point notre équation devient

$$\frac{du}{ds} = v(s, u).$$

Ceci étant, le groupe de symétries $\{g^s\}$ devient un groupe de translations le long de l'axe s, donc la *fonction v ne dépend pas de s.* Le champ de vecteurs $v(u)$

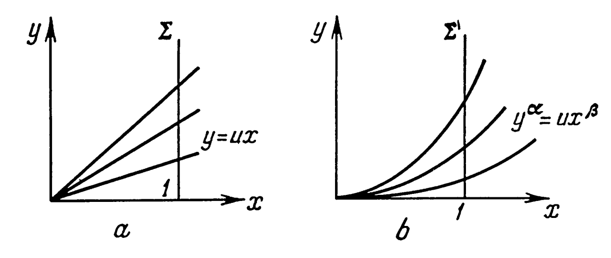


Fig. 8

sur Σ définit un champ de directions sur cette surface de $n - 1$ dimensions; si l'on connaît les courbes intégrales de ce champ, on trouve (moyennant une quadrature) les solutions de l'équation $\frac{du}{ds} = v(u)$ et, partant, les courbes intégrales de l'équation initiale.

Dans le cas particulier $n = 2$, l'introduction des coordonnées (s, u) nous conduit immédiatement à l'équation à intégrer $du/ds = v(u)$.

Remarque. En pratique on a souvent intérêt à remplacer la coordonnée s par une fonction appropriée $z(s)$. Dans un tel système de coordonnées (l'équation admettant $\{g^s\}$ pour groupe de symétries deviendra

$$\frac{du}{dz} = v(u) f(z)$$

(avec $n = 2$, on aura une équation à variables séparables).

Exemple. L'équation homogène se ramène à une équation à variables séparables aussi bien dans un système de coordonnées polaire que dans le système de coordonnées $u = y/x$, $z = x$ (fig. 8 *a*).

$\{g^s\}$ est ici le groupe à un paramètre d'homothéties de rapport e^s; dans le système de coordonnées polaires, Σ est le cercle $x^2 + y^2 = 1$, dans le second système, la droite $x = 1$; $z = e^s$.

Exercice. Trouver les coordonnées dans lesquelles l'équation quasi homogène

$$\frac{dy}{dx} = v(x, y)$$

s'intègre explicitement; α est le poids de x, β celui de y (donc v est une fonction quasi homogène de degré $\beta - \alpha$).

Solution. On peut prendre $u = y^\alpha/x^\beta$, $z = x$ (dans le domaine où $x \neq 0$). Cf. figure 8 *b*.

Exercice. Expliciter l'équation à variables séparables à laquelle se réduit l'équation de l'exercice précédent dans les coordonnées (u, z).

Solution. $y^\alpha = ux^\beta$, donc $\alpha y^{\alpha-1}\,dy = x^\beta\,du + \beta\,ux^{\beta-1}\,dx$. Si $dy = v\,dx$, alors $\alpha y^{\alpha-1} v\,dx = x^\beta du + \beta u x^{\beta-1} dx$, i.e.

$$\frac{du}{dx} = \frac{\alpha y^{\alpha-1} v - \beta\, u x^{\beta-1}}{x^\beta}.$$

Or $v(x, y) = x^{(\beta/\alpha)-1} w(u)$, donc

$$\frac{du}{dx} = \frac{\alpha w(u) - \beta u}{x}.$$

§ 2. Eclatement des singularités des équations différentielles

On décrit brièvement une importante méthode appelée résolution des singularités, éclatement ou σ-processus.

A. σ-processus

Au voisinage de tout point non singulier tous les champs de vecteurs sont de même structure simple.

Pour étudier les moindres détails des objets mathématiques au voisinage des points singuliers on a mis au point un appareil spécial dit *éclatement des singularités* qui, tel un microscope, possède un grand pouvoir de résolution. Analytiquement, cela revient à choisir au voisinage d'un point singulier des systèmes de coordonnées dans lesquels aux petits déplacements sont associées d'importantes variations des coordonnées aux voisinages des singularités.

Les coordonnées polaires jouissent de cette propriété, mais seules des fonctions transcendantes (trigonométriques) permettent d'y passer; d'un point de vue algébrique on aura intérêt à se servir d'un autre procédé appelé σ-processus ou encore éclatement des singularités.

Nous aurons besoin d'une construction auxiliaire. Soit $p : \mathbb{R}^2 \setminus O \to \mathbb{R}P^1$ un fibré standard définissant une droite projective. [Une droite projective est une variété dont les points sont des droites du plan issues de l'origine des coordonnées. L'application p associe à tout point du plan la droite reliant ce point à l'origine des coordonnées.]

Considérons le *graphe* Γ de l'application p. Ce graphe est une surface différentiable dans le produit cartésien $(\mathbb{R}^2 \setminus O) \times \mathbb{R}P^1$ (fig. 9). En plongeant $\mathbb{R}^2 \setminus O$ dans $\mathbb{R}^2$ on peut regarder le graphe comme une surface différentiable Γ dans le produit cartésien $\mathbb{R}^2 \times \mathbb{R}P^1$. La projection naturelle $\pi_1 : \mathbb{R}^2 \times \mathbb{R}P^1 \to \mathbb{R}^2$ est un difféomorphisme

de Γ sur $\mathbb{R}^2 \setminus O$. [Pour avoir une vision nette de tout cela on remarquera que Γ possède localement la forme d'un escalier spiroïdal : la droite projective est globalement difféomorphe au cercle S^1, quant au produit $\mathbb{R}^2 \times \mathbb{R}P^1$ il est difféomorphe à l'intérieur du tore.]

T h é o r è m e. *Dans* $\mathbb{R}^2 \times \mathbb{R}P^1$ *l'adhérence du graphe* Γ *de l'application* p *est une surface différentiable* $\Gamma_1 = \Gamma \cup (O \times \mathbb{R}P^1)$. *La surface* Γ_1 *est difféomorphe au ruban de Möbius* (fig. 10).

◀ Soient (x, y) des coordonnées dans le plan, $u = y/x$ une coordonnée locale affine dans $\mathbb{R}P^1$. Alors (x, y, u) est un système de coordonnées locales dans $\mathbb{R}^2 \times \mathbb{R}P^1$. Dans ce système, Γ a pour équation

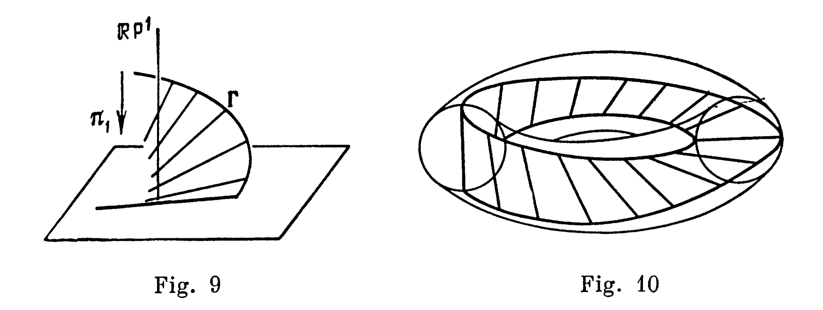


Fig. 9 Fig. 10

$y = ux$, $x \neq 0$, donc Γ_1 y aura pour équation locale $y = ux$. Cette surface est différentiable ; on l'obtient en ajoutant à la région de Γ couverte par notre système de coordonnées, la partie de la droite projective $O \times \mathbb{R}P^1$ couverte également par ce système.

On achève de démontrer que Γ_1 est différentiable en envisageant le deuxième système de coordonnées locale (x, y, v), où $x = vy$.

La projection $\pi_2 \colon \mathbb{R}^2 \times \mathbb{R}P^1 \to \mathbb{R}P^1$ fibre Γ_1 en droites. Si l'on décrit le cercle $\mathbb{R}P^1$ la droite correspondante de $\mathbb{R}^2$ tourne d'un angle π ; d'où il suit que Γ_1 est un ruban de Möbius. ▶

D é f i n i t i o n. Le passage de $\mathbb{R}^2$ à Γ_1 s'appelle *σ-processus de centre O*, ou encore *éclatement du point O* en la droite $O \times \mathbb{R}P^1$. L'application $\pi_1 \colon \Gamma_1 \to \mathbb{R}^2$ s'appelle *antisigma-processus*.

La restriction de l'application $\pi_1 \colon \Gamma_1 \to \mathbb{R}^2$ à Γ est un difféomorphisme sur $\mathbb{R}^2 \setminus O$. Donc, dans un plan, les objets géométriques présentant des singularités au point O se transposent à Γ_1. En outre, ces singularités peuvent devenir plus simples ou même disparaître.

E x e m p l e. Soient trois droites concourantes en O. Leurs images sur Γ_1 sont trois droites coupant $\mathbb{R}P^1$ en des points distincts (fig. 11).

E x e r c i c e. Soient deux courbes possédant en O un contact d'ordre n (par exemple, $n = 2$ pour $y = 0$ et $y = x^2$). Montrer que leurs images sur Γ_1 sont des courbes possédant un contact d'ordre $n - 1$ en O_1 image de O (fig. 12).

Si le σ-processus ne ramène pas les singularités à des intersections transversales, on peut l'appliquer encore une fois aux points singuliers obtenus et ainsi de suite tant qu'on n'a pas abouti à des intersections transversales. On démontre

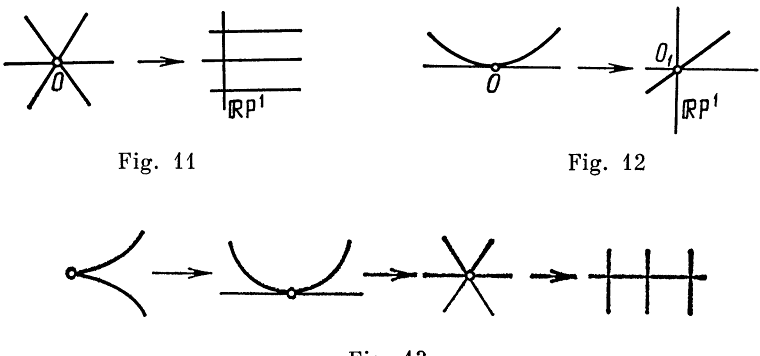


Fig. 11

Fig. 12

Fig. 13

que les singularités de toute courbe algébrique sont susceptibles d'être résolues (ramenées à des intersections transversales) en un nombre fini de pas.

E x e r c i c e. Résoudre la singularité de la courbe $x^2 = y^3$.

R é p o n s e. Cf. figure 13.

B. Formules de résolution.

En pratique, le σ-processus signifie le passage des coordonnées (x, y) aux coordonnées $(x, u = y/x)$ là où $x \neq 0$ et aux coordonnées $(v = x/y, y)$ là où $y \neq 0$ (fig. 14). Voyons ce qu'il advient alors

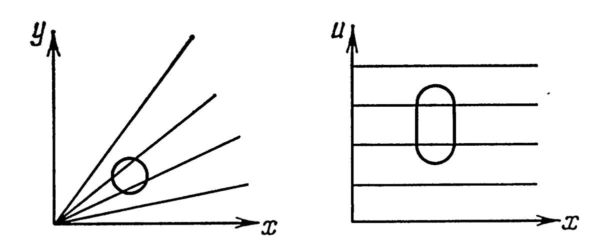


Fig. 14

de l'équation différentielle définie par un champ de vecteurs sur le plan (x, y). On suppose que le point O est un point singulier de ce champ de vecteurs.

T h é o r è m e. *Le σ-processus transforme un champ de vecteurs w présentant une singularité en O en un champ de vecteurs sur Γ prolongeable à un champ différentiable sur Γ_1.*

◀ Soit w champ de vecteurs définissant le système $\dot{x} = P(x, y)$, $\dot{y} = Q(x, y)$. Dans les coordonnées $(x, u = y/x)$ on trouve

$$\dot{x} = P(x, ux) \quad \dot{u} = (Q(x, ux) - uP(x, ux))/x.$$

Les seconds membres sont différentiables, puisque $P(0, 0) = Q(0, 0) = 0$. Dans le deuxième système de coordonnées ($v = x/y$, y) on obtient aussi un champ différentiable. ▶

R e m a r q u e. Il est possible que le champ de vecteurs obtenu s'annule sur la droite $\mathbb{R}P^1$ introduite par le σ-processus. Dans ce cas on peut diviser le champ par x dans le domaine du premier système de coordonnées et par y dans celui du second. La division ne modifie pas la direction des vecteurs du champ. On obtient donc sur Γ_1 un champ de directions, dont les points singuliers sont situés sur la droite $\mathbb{R}P^1$ mais ne la couvrent pas entièrement. Au voisinage de chaque point singulier le champ de directions est défini par un champ de vecteurs différentiable.

A toute « direction d'entrée » des courbes du champ initial en O correspond un point singulier du champ obtenu, situé sur la droite $\mathbb{R}P^1$.

Si les points singuliers O_i ne sont pas d'une structure assez simple, on peut leur appliquer le σ-processus. En poursuivant cette procédure on aboutit enfin au cas où l'une au moins des valeurs propres du linéarisé du champ en chaque point singulier est différente de zéro.

Dans de nombreux cas dès le premier σ-processus on peut se faire une idée sur le comportement des trajectoires ou des courbes intégrales au voisinage d'un point singulier. Ainsi les courbes intégrales de l'équation homogène se transforment par le changement de coordonnées $(x, y) \mapsto (x, u = y/x)$ en courbes intégrales d'équations à variables séparables.

C. Exemple. Etude du pendule en présence de frottement.

Illustrons cette méthode sur l'exemple trivial de l'équation linéaire. L'équation d'un pendule de coefficient de frottement k s'écrit $\ddot{x} + k\dot{x} + x = 0$. Elle est équivalente au système

$$\dot{x} = y, \quad \dot{y} = -ky - x$$

de plan de phase (x, y). On est conduit à l'équation homogène

$$\frac{dy}{dx} = -k + \frac{y}{x}.$$

La théorie générale dit que par le σ-processus, i.e. dans le système de coordonnées $(x, u = y/x)$, les variables deviennent séparables. En effet, $\frac{du}{dx} = -\frac{u^2 + ku + 1}{ux}$. Si l'on pose $\ln|x| = z$, on obtient

$$\frac{du}{dz} = -k - \left(u + \frac{1}{u}\right).$$

Etudions les courbes intégrales de cette équation pour diverses valeurs du coefficient $k > 0$. Le graphe de la fonction $f = u + \frac{1}{u}$ est une hyperbole

(fig. 15). Donc, celui de la fonction $-k - f(u)$ est de la forme représentée sur la figure 16. Les courbes intégrales de l'équation $\frac{du}{dz} = -k - f(u)$ seront de la forme représentée sur la figure 17. Si l'on se replace dans le plan de phase (x, y), on obtient la figure 18.

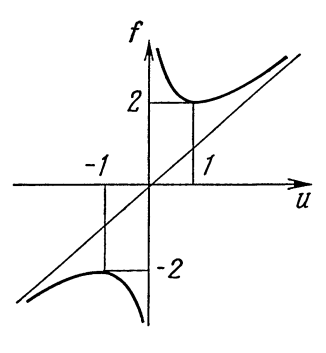


Fig. 15

Donc, pour de petites valeurs du coefficient de frottement $(0 < k < 2)$ le pendule effectue une infinité d'oscillations, pour $k \geqslant 2$ le sens du mouvement du pendule varie au plus une fois.

E x e r c i c e. Tracer les trajectoires des équations $\dot{z} = az^n$, $\dot{z} = a\bar{z}^n$, $z \in \mathbb{C}$.

D. Exemple. Période des faibles oscillations.

T h é o r è m e. *Supposons que toutes les trajectoires passant par des points voisins de la position d'équilibre O sont fermées. La limite de la période d'oscillations au voisinage du point O lorsque l'amplitude d'oscillations tend vers* 0 *est égale à la période d'oscillations dans le système linéarisé.*

◀ Le σ-processus transforme les trajectoires fermées faisant une fois le tour du point O en courbes du ruban de Möbius qui se bouclent après deux tours; faire tendre l'amplitude des oscillations vers 0 revient à faire tendre la trajectoire du ruban de Möbius vers la droite projective introduite par le σ-processus (la ligne médiane du ruban de Möbius).

Le théorème de la dépendance continue de la solution par rapport à la condition initiale nous apprend que la limite de la période des oscillations lorsque l'amplitude tend vers 0 est égale au double de la période de révolution sur la droite $\mathbb{R}P^1$ dans le système obtenu par le σ-processus. Or les vitesses du mouvement sont les mêmes sur la droite $\mathbb{R}P^1$ pour le champ donné et son linéarisé (cf. équation pour $\dot{u}$ et le point B). Il est immédiat de vérifier que toutes les trajectoires de l'équation linéarisée sont fermées. Dans le système linéaire ces trajectoires sont parcourues en des intervalles de temps égaux, puisque le champ de vecteurs linéaire reste invariant par les homothéties du plan de phase. Donc, la limite de la période d'oscillations dans le système initial est égale à la période d'oscillations dans le système linéarisé, et partant, égale tout simplement à la période des oscillations dans le système linéarisé. ▶

R e m a r q u e. La limite mentionnée s'appelle *période des faibles oscillations.*

E x e r c i c e. Calculer la période des faibles oscillations du pendule $\ddot{x} = -\sin x$ au voisinage de la position d'équilibre $x = 0$.

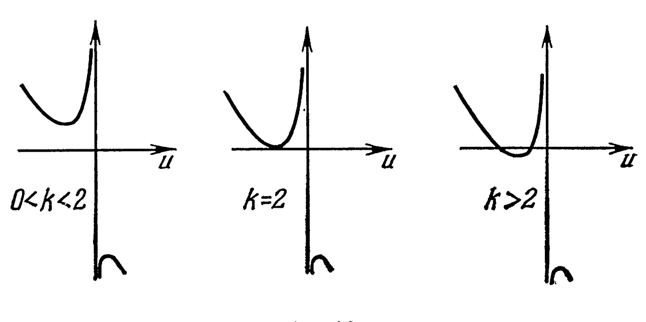


Fig. 16

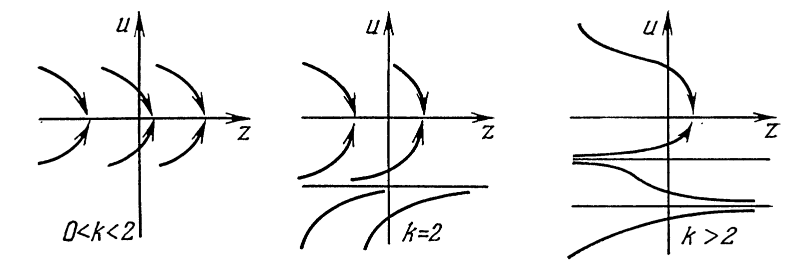


Fig. 17

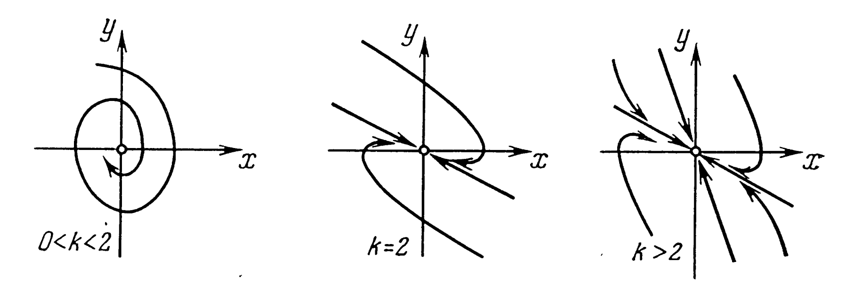


Fig. 18

§ 3. Equations non résolues par rapport aux dérivées

Dans ce paragraphe on se penche sur les principales notions de la théorie des équations différentielles, non résolues par rapport aux dérivées, du point de vue de la théorie générale des singularités des applications différentiables et de la géométrie de l'espace des jets.

A. Définitions fondamentales.

L'équation mentionnée est

$$F(x, y, p) = 0, \tag{1}$$

où $p = dy/dx$.

E x e m p l e s. 1) $p^2 = x$; 2) $p^2 = y$; 3) $y = px + p^2$.

L'espace à trois dimensions muni des coordonnées (x, y, p) s'appelle *espace des 1-jets des fonctions $y(x)$*. Deux fonctions différentiables *) y_1, y_2 admettent un même k-jet au point x_0 si $|y_1(x) -$

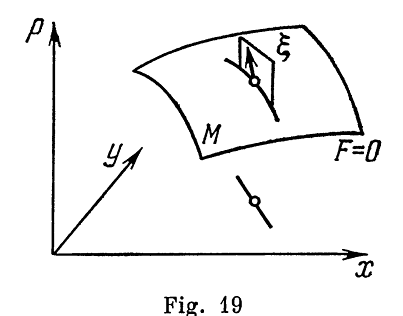


Fig. 19

$- y_2(x)| = o(|x - x_0|^k)$. Donc, un 1-jet de fonction $y(x)$ est défini par la donnée du point x, de la valeur de y en ce point et de la valeur de p en ce point.

L'équation (1) définit une surface dans l'espace des jets. Il est possible de munir cette surface d'un champ de directions en procédant comme suit. Prenons un point quelconque de l'espace des jets. Notons $dx(\xi)$, $dy(\xi)$, $dp(\xi)$ les composantes d'un vecteur ξ appliqué en ce point. Donc, dx, dy et dp ne sont pas des infiniment petits hypothétiques, mais des fonctions linéaires de ξ parfaitement définies.

Considérons au point (x, y, p) de l'espace des jets le plan composé des vecteurs ξ tels que $dy = p\,dx$. En d'autres termes, un vecteur ξ d'origine (x, y, p) est contenu dans ce plan (fig. 19) si la pente

*) On dira différentiable pour indéfiniment différentiable

de sa projection sur le plan euclidien (x, y) est égale à p. Le plan construit s'appelle *plan de contact*. Donc, en tout point de l'espace des 1-jets est appliqué un plan de contact; ensemble, ces plans forment un champ de plans de contact (ou encore une structure de contact) dans l'espace des 1-jets.

Exercice*. Existe-t-il dans l'espace des 1-jets une surface tangente en chacun de ses points au plan de contact appliqué en ce point?

Réponse. Non.

Supposons que la surface définie par l'équation (1) dans l'espace des 1-jets est différentiable. [Cette restriction n'est pas très importante, puisque pour une fonction lisse générique F la valeur 0 n'est pas critique et l'ensemble de niveau 0 est différentiable; si tel n'est pas le cas pour la fonction envisagée, on peut, en faisant subir à cette dernière un changement aussi petit que l'on veut, rendre différentiable l'ensemble de niveau 0: par exemple, en lui ajoutant une petite constante (cf. théorème de Sard, § 10, point F).]

Considérons un point de la surface différentiable M définie par l'équation (1) et supposons que le plan tangent à cette surface en ce point est distinct du plan de contact. Ces plans se coupent alors suivant une droite. De plus, les plans tangents et les plans de contact en tous les points voisins se coupent suivant des droites, si bien qu'on obtient un champ de directions au voisinage du point envisagé.

On appelle *courbes intégrales* de l'équation (1) les courbes intégrales du champ de directions obtenu sur la surface M. Intégrer (ou étudier) l'équation (1) revient à trouver (ou étudier) ces courbes. Le lien entre les courbes intégrales de M et les graphes des solutions de l'équation (1) est étudié plus bas. A noter que les courbes intégrales sur M sont définies non pas en termes de solutions de l'équation (1), mais en termes de plans de contact.

B. Points réguliers et courbe discriminant.

Appelons *verticale* la direction de l'axe p dans l'espace des jets. Soit M surface différentiable d'équation (1) dans l'espace des jets. Considérons la projection suivant la verticale:

$$\pi: M \to \mathbb{R}^2, \quad \pi(x, y, p) = (x, y).$$

Définition. Un point de M est *régulier* s'il n'est pas point critique de l'application π.

Autrement dit, un point de M est régulier si le plan tangent en ce point n'est pas vertical ou encore si l'application π sur le plan (x, y) est un difféomorphisme au voisinage de ce point.

L'ensemble des valeurs critiques de l'application (c'est-à-dire la projection de l'ensemble des points critiques) s'appelle *courbe discriminant* de l'équation (1).

E x e m p l e. L'axe des y est courbe discriminant de l'équation $p^2 = x$, l'axe des x, de l'équation $p^2 = y$ (fig. 20).

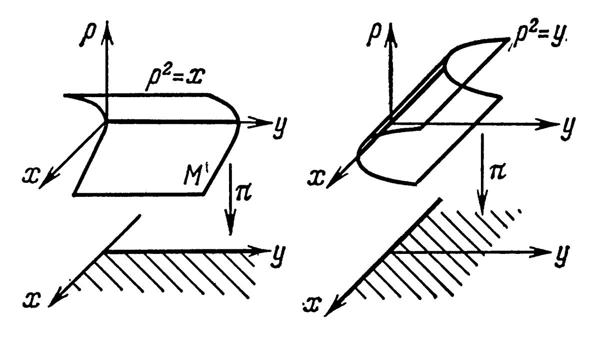


Fig. 20

Soit un point régulier de M. Le théorème des fonctions implicites nous dit qu'au voisinage de ce point la surface M est le graphe d'une fonction différentiable $p = v(x, y)$.

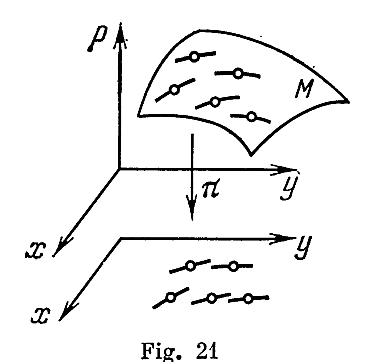


Fig. 21

T h é o r è m e. *La projection sur le plan* (x, y) *associe aux courbes intégrales de l'équation* (1) *sur* M *au voisinage d'un point régulier les courbes intégrales de l'équation*

$$\frac{dy}{dx} = v(x, y)$$

au voisinage de la projection de ce point (fig. 21).

◀ La projection sur le plan (x, y) du plan de contact est, par définition de ce dernier, une droite du champ de direction (2). Donc les courbes intégrales se correspondent. ▶

R e m a r q u e. Les projections des courbes intégrales globales de l'équation (1) sur le plan (x, y) ne sont généralement courbes intégrales d'aucun champ de directions. Les projections des courbes

intégrales de l'équation (1) sur le plan (x, y) possèdent dans le cas général des points de rebroussement sur la courbe discriminant, cependant pour certaines équations (1) elles restent différentiables en les points de la courbe discriminant.

C. Exemples.

Exemple 1. $p^2 = x$ (fig. 22).

La surface M est un cylindre parabolique. L'axe des y est courbe discriminant. Pour déterminer les courbes intégrales on a intérêt à prendre pour coordon-

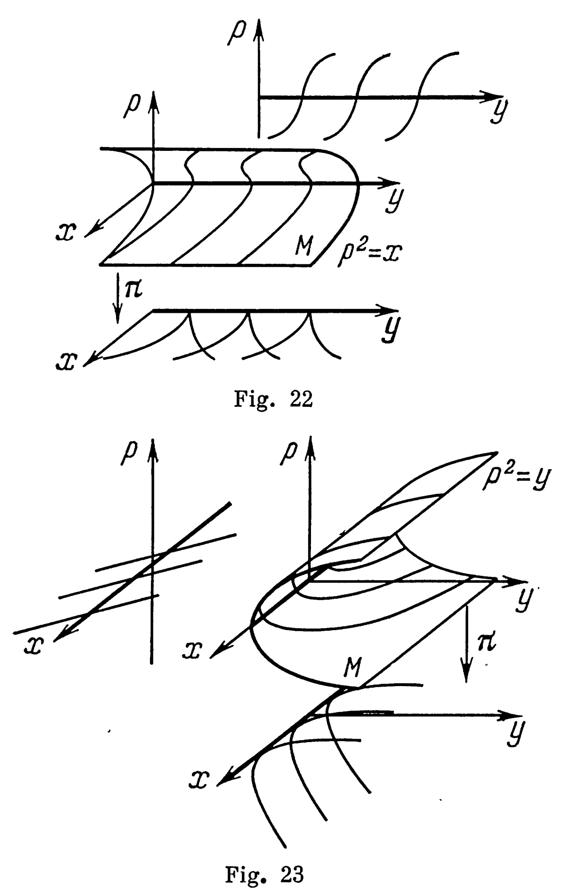


Fig. 22

Fig. 23

nées sur M non pas x et y mais p et y (d'autant plus que ces dernières sont globales).

Entre les composantes dx, dy, dp d'un vecteur ξ d'origine $(x, y, p) \in M$ et appartenant à notre champ de directions, on a les relations suivantes:

$$\begin{cases} p^2 = x \text{ (condition d'appartenance à } M); \\ 2p\,dp = dx \text{ (condition de tangence à } M); \\ dy = p\,dx \text{ (condition d'appartenance au plan de contact).} \end{cases}$$

Donc, dans les coordonnées (p, y) les courbes intégrales sont celles de l'équation $dy = 2p^2\,dp$.

Ainsi, les courbes intégrales ont pour équation $y + C = \frac{2}{3} p^3$, $x = p^2$. Leurs projections sur le plan (x, y) sont des paraboles semi-cubiques.

Exemple 2. $p^2 = y$ (fig. 23).

En procédant comme dans l'exemple précédent, on obtient

$$\begin{cases} p^2 = y, \\ 2p\,dp = dy, \\ dy = p\,dx. \end{cases}$$

Si l'on se place dans les coordonnées x et p, on obtient $p\,(dx - 2dp) = 0$, d'où soit $p = 0$, $y = 0$, soit

$$x = 2p + C,\ y = p^2.$$

Les projections de ces courbes sur le plan (x, y) sont des paraboles tangentes à la courbe discriminant $y = 0$.

Exemple 3. (Equation de Clairault) $y = px + f(p)$ (fig. 24).

La surface M est réglée (car ses traces sur les plans $p =$ const sont des droites). Pour coordonnées sur M il est commode de prendre

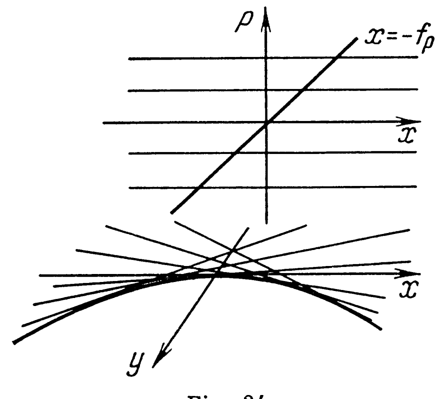


Fig. 24

x et p. Les courbes intégrales se déterminent à partir des relations

$$\begin{cases} y = px + f(p), \\ dy = p\,dx + x\,dp + f'\,dp, \\ dy = p\,dx. \end{cases}$$

On trouve $(x + f')\,dp = 0$. Les points en lesquels $x + f' = 0$ sont critiques, les autres, réguliers. Sur le plan de coordonnées (x, p), les courbes intégrales sont des droites $p =$ const $= C$ qui, en général, coupent la ligne des points critiques $(x + f' = 0)$.

Les projections des courbes intégrales sur le plan (x, y) sont des droites $y = Cx + f(C)$ tangentes à la courbe discriminant.

[En toute rigueur, les points d'intersection avec la ligne critique n'appartiennent pas aux courbes intégrales, car le champ de directions n'y est pas défini : le plan de contact y est tangent à M.]

La courbe discriminant se déduit à partir des relations

$$y = px + f(p), \quad x + f' = 0.$$

Si, par exemple, $f(p) = -p^2/2$, la courbe discriminant est alors la parabole $y = x^2/2$ et les projections des courbes intégrales, des tangentes. La théorie de l'équation de Clairault est rattachée à d'importantes notions mathématiques : la transformation de Legendre et la dualité projective.

D. Transformation de Legendre.

Soit donnée une fonction $f(x)$. On appelle *transformée de Legendre* *) de $f(x)$ une fonction g d'une nouvelle variable p, définie

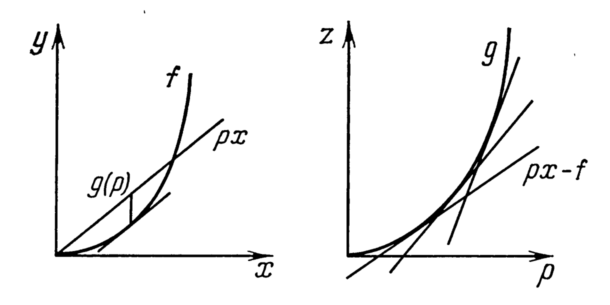


Fig. 25

comme suit. Considérons le graphe de f sur le plan (x, y). Traçons la droite $y = px$. Trouvons le point du graphe le plus éloigné de la droite par rapport à la direction de l'axe des y. Considérons la différence des ordonnées des points de la droite et de la courbe. Cette différence est par définition la valeur de la fonction g au point p (fig. 25) :

$$g(p) = \sup_x (px - f(x)).$$

Exemple 1. Soit $f(x) = x^2/2$. La transformée de Legendre est $g(p) = p^2/2$.

Exemple 2. Soit $f(x) = x^\alpha/\alpha$. On a $g(p) = p^\beta/\beta$, où $\alpha^{-1} + \beta^{-1} = 1$ (x, p, α et β sont non négatifs).

Si la fonction f est strictement convexe ($f'' > 0$) et sa dérivée définit un difféomorphisme d'une droite sur une droite, alors la fonction g est aussi strictement convexe ; et, de plus, le suprémum est réalisé en un seul point x, celui où $f'(x) = p$. Les points tels

*) Les transformations de Legendre désignent plusieurs objets mathématiques, rattachés aussi aux noms de Minkowski et de Young. Notre but n'est pas de consacrer, par pédanterie, cette terminologie.

que x et p sont dits *associés* dans la transformation de Legendre.

Théorème. *On a l'inégalité*

$$f(x) + g(p) \geqslant px.$$

Si f est strictement convexe et f' un difféomorphisme sur, alors une condition nécessaire et suffisante pour que l'égalité soit réalisée est que les points x et p soient associés.

◀ La fonction $px - f(x)$ est $\leqslant$ à son suprémum. ▶

Exemple. On a $px \leqslant x^\alpha/\alpha + p^\beta/\beta$, $\forall x < 0$, $\forall p < 0$.

Dans les corollaires formulés plus bas on suppose que les fonctions f et g sont strictement convexes et leurs dérivées définissent des difféomorphismes d'une droite sur une droite.

Corollaire. *La transformation de Legendre est involutive: la transformée de Legendre de la fonction $g(p)$ est la fonction $f(x)$ (sous réserve que les coordonnées soient convenablement désignées).*

◀ En effet, l'inégalité du théorème précédent est symétrique par rapport à f et g. ▶

Corollaire. *On passe par une transformation de Legendre d'une fonction strictement convexe g définissant l'équation de Clairault $y = px - g(p)$ à la fonction f définissant l'enveloppe des solutions avec la formule $y = f(x)$.*

◀ Le graphe de la fonction f est l'enveloppe de ses tangentes. ▶

Remarque. La transformée de Legendre d'une fonction de n variables se définit de façon analogue et possède les mêmes propriétés. Si $x \in \mathbb{R}^n$ alors p est un point de l'espace vectoriel dual (l'espace $\mathbb{R}^{n*}$ des fonctions linéaires sur $\mathbb{R}^n$).

E. Dualité projective.

La transformation de Legendre est un cas particulier d'une construction plus générale de géométrie projective. Soit un espace projectif de dimension n, noté $\mathbb{R}P^n$.

Un point de l'espace projectif est défini par un vecteur non nul x de l'espace affine $\mathbb{R}^{n+1}$, à un facteur multiplicatif près non nul. Ce qu'on note brièvement

$$\mathbb{R}P^n = (\mathbb{R}^{n+1} \setminus O)/(\mathbb{R} \setminus O).$$

Un hyperplan d'un espace projectif est constitué de tous les points de cet espace projectif tels que les points associés de l'espace affine appartiennent à un même hyperplan passant par zéro.

Considérons l'ensemble de tous les hyperplans d'un espace projectif de dimension n. Cet ensemble est lui-même un espace projectif de dimension n.

En effet, un hyperplan d'un espace projectif est défini par l'équation homogène

$$(a, x) = 0, \quad x \in \mathbb{R}^{n+1}, \quad a \in \mathbb{R}^{n+1*} \setminus 0,$$

où $\mathbb{R}^{n+1*}$ est l'espace des fonctions linéaires de $\mathbb{R}^{n+1}$ (cet espace qui est vectoriel et de dimension $n+1$ est dit dual de l'espace $\mathbb{R}^{n+1}$).

Donc, à un hyperplan d'un espace projectif est associé un vecteur non nul de $\mathbb{R}^{n+1*}$ défini à un facteur multiplicatif près non nul. Par suite, l'*ensemble de tous les hyperplans de* $\mathbb{R}P^n$ *possède une structure naturelle d'espace projectif de dimension* n:

$$\mathbb{R}P^{n*} = (\mathbb{R}^{n+1*} \diagdown O) / (\mathbb{R} \diagdown O).$$

L'espace projectif des hyperplans de l'espace projectif donné $\mathbb{R}P^n$ est dit *dual de* $\mathbb{R}P^n$ et se note $\mathbb{R}P^{n*}$. Par exemple, l'espace de toutes les droites d'un plan projectif est un plan projectif dual du plan initial.

On notera que la dualité est une notion réciproque, c'est-à-dire $\mathbb{R}P^{n**} = \mathbb{R}P^n$. Ceci découle de la symétrie de a et x dans l'équation de l'hyperplan $(a, x) = 0$.

Exemples. Les droites passant par un point d'un plan projectif forment, comme il est aisé de le constater, une droite sur le plan dual. Les droites concourantes comprises dans un angle ayant le point de concours pour sommet forment un segment sur le plan dual.

Les tangentes à une conique non dégénérée d'un plan projectif forment une conique non dégénérée sur le plan dual. D'une façon générale, les tangentes à une courbe différentiable quelconque forment une courbe (pas forcément différentiable) sur le plan dual, appelée courbe duale de l'initiale.

Théorème. *Les graphes d'une fonction strictement convexe et de sa transformée de Legendre sont projectivement duaux.*

◀ Considérons toutes les droites non parallèles à l'axe des y d'un plan affine rapporté à un système de coordonnées (x, y). Ces droites forment un plan : on peut définir une droite à l'aide de l'équation $y = px - z$ et considérer (p, z) comme les coordonnées affines du nouveau plan. Dans ce cas la transformation de Legendre consiste à passer du graphe de la fonction f à une famille de tangentes à cette courbe: lorsqu'un point décrit le graphe de la fonction f sur le plan (x, y), la tangente au graphe de la fonction f décrit sur le plan (p, z) la courbe représentative de la transformée de Legendre $z = g(p)$. ▶

Donc, *la transformation de Legendre consiste simplement à passer de la courbe à sa duale projective en coordonnées affines.*

Exemple. Supposons que le graphe d'une fonction f est une ligne polygonale convexe (vers le bas). On appelle *droite d'appui* une droite située au-dessous du graphe et ayant un point commun avec lui *). Considérons toutes les droites d'appui de la ligne polygonale.

---

*) D'une façon générale, un *hyperplan d'appui* d'un corps convexe est un plan possédant un point commun avec ce corps et tel que ce corps soit entièrement compris dans l'un des demi-espaces en lesquels ce plan partage l'espace.

Il est immédiat de vérifier qu'elles forment une ligne polygonale convexe sur le plan dual. En effet, les droites d'appui recouvrent l'angle de deux segments de la ligne polygonale initiale, donc elles forment un segment sur le plan dual. De façon analogue, les sommets de la ligne duale sont les images des segments de la ligne de départ.

Avec la dualité projective on peut étudier des cas plus généraux qu'avec la transformation de Legendre.

Exercice. Tracer la courbe projectivement duale de la courbe représentée sur la figure 26.

Indication. Les images des tangentes doubles à la courbe de départ sont des points de self-intersection de la courbe duale, donc la courbe duale possède 4 points de self-intersection.

Les images des points d'inflexion de la courbe de départ sont des points de rebroussement de la courbe duale. En effet, si $f = x^3$, la tangente au graphe

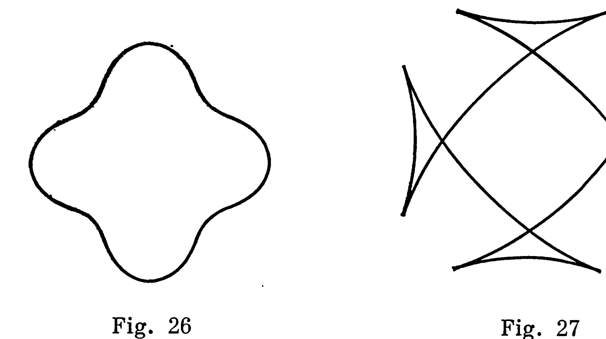


Fig. 26 Fig. 27

au point $x = t$ a pour équation $p = 3t^2$, $z = 2t^3$. Ce graphe présente un point de rebroussement dans le plan rapporté aux coordonnées (p, z). Ainsi, la courbe duale possède 8 points de rebroussement, disposés deux par deux entre les points de self-intersection.

Supposons que la courbe initiale est constituée de deux ellipses légèrement arrondies à leurs points d'intersection (on ne considère pas les portions d'ellipses inférieures l'une à l'autre).

La courbe duale comporte également un couple d'ellipses. Les images des points d'intersection des ellipses sont des tangentes doubles aux ellipses duales. Il est aisé de voir maintenant comment la courbe duale se déduit du couple d'ellipses concourantes avec leurs tangentes doubles (fig. 27).

F. Transformation de Legendre et normes conjuguées.

Définition. On appelle *norme* dans $\mathbb{R}^n$ une fonction réelle, non négative, convexe, positivement homogène, du premier degré, paire, nulle seulement en l'origine des coordonnées:

$$f \geqslant 0,\ f(x) = 0 \Leftrightarrow x = 0,\ f(\lambda x) = |\lambda| f(x),$$
$$f(x + y) \leqslant f(x) + f(y).$$

Soit f norme dans $\mathbb{R}^n$. La fonction f est définie par un ensemble sur lequel elle est égale à l'unité. Cet ensemble est une hypersurface de $\mathbb{R}^n$ convexe, à symétrie centrale par rapport au zéro. Inversement, tout corps de $\mathbb{R}^n$ compact, convexe, symétrique par rapport au zéro, comprenant le zéro de $\mathbb{R}^n$ à son intérieur, définit une norme unique égale à un sur son bord. Cette hypersurface $f = 1$ s'appelle *sphère unité* de norme f.

Exercice 1. Dans $\mathbb{R}^3$ trouver les sphères unités de normes:

$$\text{a) } f=\sqrt{(x, x)},\ \text{b) } f=\max|x_i|,\ \text{c) } f=\sum |x_i|.$$

Considérons l'espace $\mathbb{R}^{n*}$ conjugué de $\mathbb{R}^n$.

Définition. La *norme conjuguée* dans $\mathbb{R}^{n*}$ est définie par

$$g(p)=\max_{f(x)\leqslant 1} |(p, x)|.$$

Il est immédiat de vérifier que g est une norme réelle.

La relation de conjugaison est réciproque, puisque l'inégalité servant à sa définition est symétrique: $|(p, x)| \leqslant f(x)\, g(p)$.

Associons à tout point p de l'espace conjugué l'hyperplan $p = 1$ de l'espace de départ.

Théorème. *La sphère unité de la norme conjuguée est l'ensemble des hyperplans d'appui de la sphère unité de la norme initiale.*

◀ La condition $g(p) = 1$ signifie que (p, x) possède un maximum égal à un sur la boule initiale, c'est-à-dire le plan $p = 1$ est un plan d'appui pour la sphère initiale. ▶

L'ensemble de tous les hyperplans d'appui d'une hypersurface donnée s'appelle *hypersurface duale convexe*. Donc, les sphères unités des normes duales sont duales.

Exercice 2. Dans $\mathbb{R}^3$ trouver les surfaces duales à a) une sphère; b) un tétraèdre; c) un cube; d) un octaèdre.

Exercice 3. Trouver les normes conjuguées à celles de l'exercice 1.

De ce qui précède il résulte qu'on passe d'une hypersurface convexe à sa duale localement par une transformation de Legendre.

G. Problème des enveloppes d'une famille de courbes planes.

Deux fonctions différentiables de deux variables

$$x = x(s, t),\ y = y(s, t)$$

définissent une famille de courbes sur le plan, paramétrisées par le paramètre t (qui définit un point sur la courbe) et numérotées par le paramètre s (qui indique le numéro de la courbe).

Exercice. Tracer les familles de courbes

a) $x = (s+t)^2$, $y = t$,

b) $x = s + st + t^3$, $y = t^2$ (s et t sont petits);

c) $x = (s+t^2)^2$, $y = t$.

Réponse. Cf. figure 28.

On montre que l'enveloppe d'une famille générique est une courbe dont les seuls points singuliers sont des points de rebroussement (comme la parabole semi-cubique) et des points de self-intersection; de plus, au voisinage de chaque point régulier de l'enveloppe la famille se réduit à l'une des formes normales a), b) et c) par un changement différentiable des coordonnées $X(x, y)$, $Y(x, y)$ et des paramètres $S(s)$, $T(s, t)$.

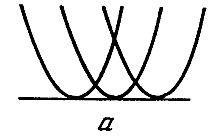
Par ailleurs, au voisinage d'un point générique de la courbe discriminant l'équation non résolue par rapport à la dérivée se ramène généralement à la forme normale

$$\left(\frac{dy}{dx}\right)^2 = x$$

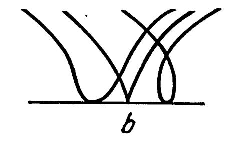
par un changement différentiable des coordonnées $X(x, y)$, $Y(x, y)$ (cf. § 4).

Dans ces coordonnées les projections des courbes intégrales sur le plan (X, Y) sont des paraboles semi-cubiques. Donc, la courbe discriminant est l'enveloppe des projections des courbes intégrales seulement pour des équations exception-

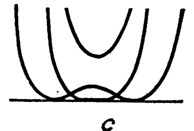
nelles (par exemple, pour l'équation de Clairault). En particulier, si l'on modifie



a



b



c

Fig. 28

légèrement l'équation de Clairault, on constate que la courbe discriminant d'enveloppe se transforme en lieu géométrique des points de rebroussement des projections des courbes intégrales.

§ 4. Forme normale d'une équation résolue par rapport à la dérivée au voisinage d'un point singulier régulier

On étudie les singularités de la famille de courbes intégrales d'une équation différentielle générique non résolue par rapport à la dérivée.

A. Points singuliers.
Soit l'équation

$$F(x, y, p) = 0, \quad p = dy/dx, \tag{1}$$

définie par une fonction différentiable F dans un domaine.

On admettra que l'équation (1) définit une surface différentiable dans un espace de jets à trois dimensions, muni des coordonnées (x, y, p). Pour cela il suffit d'après le théorème des fonctions implicites qu'aux zéros de F ne s'annule pas la différentielle totale de F; ce que nous supposerons.

Considérons la projection de la surface $F = 0$ sur le plan coordonnée (x, y) parallèlement à la p-direction.

Définition. Un point de la surface $F = 0$ est *singulier pour l'équation* (1) si la projection $(x, y, p) \mapsto (x, y)$ n'est pas au voisinage de ce point un difféomorphisme local de cette surface sur le plan (x, y).

Le théorème des fonctions implicites nous dit que les points singuliers sont les points de la surface $F = 0$ en lesquels $\partial F/\partial p = 0$.

B. Courbe criminant
Considérons l'ensemble de tous les points singuliers de l'équation (1). Cet ensemble est défini par deux équations, $F = 0$ et $\partial F/\partial p = 0$ dans l'espace à trois dimensions des jets. Donc les points singuliers forment « généralement » une courbe.

Définition. L'ensemble des points singuliers de l'équation $F = 0$ dans l'espace des jets (x, y, p) est appelé *courbe criminant* *) de l'équation.

Le théorème des fonctions implicites nous dit que la courbe criminant est une courbe de l'espace des jets qui est différentiable au voisinage de chacun des ses points en lequel le rang de la dérivée de l'application $(x, y, p) \mapsto (F, \partial F/\partial p)$ est maximal (égal à 2).

---

*) Ce terme a été emprunté à V. Zalgaller, Théorie des enveloppes, Naouka, 1975, page 98 (en russe). L'auteur espère qu'il sera adopté.

C. Courbe discriminant.

D é f i n i t i o n. La projection de la courbe criminant sur le plan (x, y) parallèlement à la p-direction s'appelle *courbe discriminant* *).

Le théorème des fonctions implicites nous apprend que le voisinage d'un point de la courbe criminant se projette difféomorphement sur le plan (x, y) parallèlement à la p-direction si la courbe criminant n'est pas tangente à la p-direction en ce point.

R e m a r q u e. La courbe discriminant peut même sous ces conditions être à singularités.

Celles-ci proviennent de ce qu'un point de la courbe discriminant peut être la projection, en général, de plusieurs points de la courbe criminant. Ces singularités seront « en général » des points de self-intersection de la courbe discriminant. Pour l'équation générique, la courbe discriminant sera constituée au voisinage d'un tel point de deux branches se coupant sous un angle non nul.

Les points de tangence de la courbe criminant à une p-direction sont dans «le cas général» les antécédents des points de rebroussement de la courbe discriminant.

Les singularités plus complexes de la courbe discriminant, en dehors des points de self-intersection et de rebroussement, s'éliminent par une faible variation de l'équation. Les singularités des deux types cités ne disparaissent pas par une faible déformation de l'équation, elles ne sont que légèrement déplacées.

D. Points de tangence de la courbe criminant à un plan de contact.

L'espace des jets admet en chacun de ses points (x, y, p) un *plan de contact* $dy = p\,dx$. En particulier, un tel plan existe en chaque point de la courbe criminant. La tangente en un point de la courbe criminant peut être contenue dans le plan de contact ou le couper.

D é f i n i t i o n. Un point de la courbe criminant est *point de tangence à un plan de contact* si la tangente à la courbe criminant en ce point est contenue dans ce plan de contact.

A noter que les points de tangence de la courbe criminant à p-direction sont également des points de tangence d'un plan de contact. En effet, tout plan de contact contient une p-direction en chaque point.

E. Points singuliers réguliers.

D é f i n i t i o n. Un point singulier de l'équation (1) est *régulier* si les équations de la courbe criminant vérifient les conditions du théorème des fonctions implicites en ce point, c'est-à-dire **)

$$\text{rang}\,((x, y, p) \longmapsto (F, F_p)) = 2,$$

et si la courbe criminant n'est pas tangente au plan de contact en ce point.

E x e m p l e. Soit l'équation $p^2 = x$. La courbe criminant est définie par les équations $p = 0$, $x = 0$. C'est l'axe des y. La condition sur le rang est réalisée. Le vecteur (0, 1, 0) tangent à la courbe criminant n'est pas contenu dans le plan de contact $dy = 0\,dx$. Donc, tout point singulier de l'équation $p^2 = x$ est régulier.

R e m a r q u e. Tous les points singuliers de l'équation générique sont réguliers: les points non réguliers sont répartis discrètement sur la courbe criminant. Le cas échéant on peut les répartir discrètement par une faible déformation de l'équation. (La justification de ce fait et des généralités précédentes se base sur le théorème de Sard, § 10.)

---

*) Cette définition est une reformulation de celle de la courbe discriminant du § 3.

**) On appelle rang d'une application le rang de sa dérivée.

F. Théorème de la forme normale.

Théorème. *Soit* (x_0, y_0, p_0), *point singulier régulier de l'équation* $F(x, y, p) = 0$. *Il existe un difféomorphisme du voisinage du point* (x_0, y_0) *du plan* (x, y) *sur un voisinage du point* $(0, 0)$ *du plan* (X, Y), *ramenant l'équation* $F = 0$ *à la forme* $P^2 = X$ (*où* $P = dY/dX$).

Nota. L'équation $F = 0$ définit une surface dans l'espace à trois dimensions des éléments linéaires sur le plan (x, y). Un difféomorphisme du plan associe à tout élément linéaire un nouvel élément linéaire. On affirme qu'au voisinage d'un point singulier régulier on peut appliquer une partie de la surface $F = 0$ sur une partie de la surface $P^2 = X$ au voisinage du point $(X = 0, Y = 0, P = 0)$.

Corollaire. *La famille des courbes intégrales de l'équation* (1) *est difféomorphe au voisinage d'un point singulier régulier à la famille des paraboles semi-cubiques* $y = x^{3/2} + C$.

◀ Le difféomorphisme mentionné dans le théorème associe aux courbes intégrales de l'équation (1) dans le plan (x, y) les courbes intégrales de l'équation $P^2 = X$ dans le plan (X, Y). Ces dernières sont des paraboles semi-cubiques ayant leurs « nez » sur la courbe discriminant: $dY/dX = \sqrt{X}$, $Y = \frac{2}{3} X^{3/2} + C$. ▶

G. Démonstration du théorème de la forme normale.

◀ 1°. *Réduction au cas où la courbe criminant est l'axe des* y.

Soit (x_0, y_0, p_0) point singulier régulier de l'équation $F(x, y, p) = 0$. La courbe discriminant est alors différentiable au voisinage du point (x_0, y_0). Projetons sur le plan (x, y) les plans de contact aux points de la courbe criminant. On obtient au voisinage du point (x_0, y_0) une famille différentiable de droites non tangentes à la courbe discriminant.

Choisissons maintenant un système de coordonnées locales sur le plan (x, y) au voisinage du point (x_0, y_0), tel que

1) l'équation de la courbe discriminant soit $x = 0$;

2) les lignes $y =$ const coupent la courbe discriminant dans les directions que nous venons juste de construire.

Nous désignerons comme toujours ces coordonnées par (x, y) et la dérivée dy/dx par p. Les coordonnées du point singulier (x_0, y_0, p_0) sont maintenant $(0, 0, 0)$.

2°. *Analyse des conditions de régularité.*

Dans le nouveau système de coordonnées la courbe criminant est l'axe des y; sur cet axe on a $x = 0$, $p = 0$ ($y =$ const). D'où il suit que $F(0, y, 0) = 0$, $F_p(0, y, 0) = 0$. La condition de régularité de la courbe criminant s'ecrit maintenant

$$\det \left| \frac{D(F, F_p)}{D(x, p)} \right| \neq 0, \text{ i.e. } \det \begin{vmatrix} F_x & F_{xp} \\ F_p & F_{pp} \end{vmatrix} \neq 0,$$

(puisque $F_y = 0$, $F_{yp} = 0$ aux points de la courbe criminant). D'autre part, $F_p = 0$ sur la courbe criminant. Donc, la condition de régularité de la courbe criminant s'écrit

$$F_x(0, y, 0) \neq 0, \quad F_{pp}(0, y, 0) \neq 0.$$

La condition de non-tangence à un plan de contact est automatiquement réalisée.

Développons F en série de Taylor sur p jusqu'au terme d'ordre 2:

$$F(x, y, p) = A(x, y) + pB(x, y) + p^2 C(x, y, p).$$

Des relations précédentes il résulte que $A(0, y) = 0$, $B(0, y) = 0$. Donc

$$A(x, y) = x\alpha(x, y), \quad B(x, y) = x\beta(x, y),$$

où α et β sont des fonctions différentiables.

Les conditions de régularité de la courbe criminant s'écrivent

$$A_x(0,\ y) \neq 0,\quad C\ (0,\ y,\ 0) \neq 0.$$

Nous pouvons même admettre pour la suite que $C > 0$, $A_x < 0$ (le cas échéant on change les signes de F et / ou de x). En définitive, $\alpha\ (0,\ 0) < 0$, $C\ (0,\ 0,\ 0) > > 0$.

3° *Etude du trinôme du second degré.*

Si la relation $F = 0$ est un trinôme du second degré en p de coefficient C, B, A, on obtient

$$p=\frac{-B\pm\sqrt{B^2-4AC}}{2C}=\frac{-x\beta\pm\sqrt{x\gamma}}{2C},$$

où $\gamma = -4\alpha C + x\beta^2$ est une fonction de $(x,\ y,\ p)$; en outre, $\gamma\ (0,\ 0,\ 0) = = -4\alpha\ (0,\ 0,\ 0)\ C\ (0,\ 0,\ 0) > 0$.

Supposons, enfin, que $x = \xi^2$. On obtient si l'on ne retient que le signe « $+$ » dans « $\pm$ »:

$$p=\frac{-\xi^2\beta\ (\xi^2,\ y)+\xi\sqrt{\gamma\ (\xi^2,\ y,\ p)}}{2C\ (\xi^2,\ y,\ p)}.$$

Appliquons à cette équation en $p\ (\xi,\ y)$ le théorème des fonctions implicites. On obtient la solution $p = \xi\omega\ (\xi,\ y)$, où ω est une fonction différentiable, $\omega\ (0,\ 0) \neq 0$.

4° *Equation différentielle en* $y\ (\xi)$.

On remarquera que $p = dy/dx = dy/2\xi\ d\xi$. On obtient donc l'équation différentielle

$$dy/d\xi = 2\xi^2\omega\ (\xi,\ y),\quad \omega\ (0,\ 0) \neq 0. \tag{2}$$

Les courbes intégrales du plan $(\xi,\ y)$ coupent l'axe $\xi = 0$ et présentent un contact du second ordre avec les lignes $y = \text{const}$. L'équation possède donc une intégrale première de la forme $I\ (\xi,\ y) = y - \xi^3 K\ (\xi,\ y)$, où K est une fonction différentiable, $K\ (0,\ 0) \neq 0$ (I est la coordonnée du point d'intersection avec l'axe $\xi = 0$; $K \neq 0$ puisque $\omega \neq 0$).

5° *Construction des coordonnées normalisantes.*

Décomposons K en une partie paire et une partie impaire sur ξ:

$$K\ (\xi,\ y) = L\ (\xi^2,\ y) + \xi M\ (\xi^2,\ y).$$

L et M sont des fonctions différentiables de x et y; $L\ (0,\ 0) \neq 0$. Dans ces notations $I\ (\xi,\ y) = y - \xi^4 M\ (\xi^2,\ y) - \xi^3 L\ (\xi^2,\ y)$. Soient les variables Y, Ξ

$$\Xi=\xi\sqrt[3]{L\ (\xi^2,\ y)},\qquad Y=y-\xi^4 M\ (\xi^2,\ y).$$

On a alors $I = Y - \Xi^3$.

Posons encore

$$X = \Xi^2.$$

Il vient

$$X=x\sqrt[3]{L^2\ (x,\ y)},\qquad Y=y-x^2 M\ (x,\ y).$$

Ces formules définissent un difféomorphisme local du plan au voisinage du point 0, 0), puisque $L\ (0,\ 0) \neq 0$. L'intégrale première s'écrit

$$I=Y-X^{3/2}.$$

Donc $(dY/dX)^2 = (9/4)\ X$ et la forme normale du théorème se déduit par une homothétie de rapport $k > 1$ de l'un des axes de coordonnées. ▶

H. Remarques.

1. Le point fort de la démonstration est la substitution $x = \xi^2$, c'est-à-dire le passage à un recouvrement à deux feuillets du plan (x, y) avec une ramification le long de la courbe discriminant. Pour des raisons topologiques (en réalité, dans un domaine complexe) il est clair à priori que la bivalence de $p(x, y)$ n'a plus lieu sur ce recouvrement et l'équation se décompose en deux. La considération du trinôme du second degré n'a été nécessaire que pour justifier ce fait dans le plan réel. Il ne reste plus qu'à ramener l'équation (2) obtenue sur le recouvrement à la forme normale par un difféomorphisme du recouvrement abaissable sur le plan initial, ce qu'on réalise aisément par décomposition de l'intégrale première en une composante paire et une impaire en ξ.

La première démonstration du théorème de la forme normale est due à Y. Brodsky; elle est basée sur le travail de R. Thom qui n'a réduit l'équation qu'à la forme $p^2 = xE(x, y)$.

2. Notre démonstration utilise la représentation d'une fonction paire par une fonction du carré de l'argument. Evidente pour les fonctions analytiques (ou pour les séries formelles) cette représentation doit être justifiée pour les fonctions indéfiniment différentiables.

En effet, une fonction paire indéfiniment différentiable peut être considérée comme une fonction du carré de l'argument définie sur le demi-axe positif. Elle est indéfiniment différentiable en tous les points de ce demi-axe y compris au point zéro. On demande de la représenter comme la restriction au demi-axe positif d'une fonction indéfiniment différentiable sur l'axe tout entier.

L'existence d'une telle représentation signifie que la fonction admet un prolongement indéfiniment différentiable sur le demi-axe positif. Cette représentation est affirmée par un théorème (de E. Borel) de l'existence de fonctions indéfiniment différentiables sur la droite possédant une série de Taylor quelconque à l'origine des coordonnées. Nous omettrons la démonstration de ce théorème, qui, d'ailleurs, n'est pas compliquée.

§ 5. Equation stationnaire de Schrödinger

Dans ce paragraphe on expose les notions mathématiques élémentaires de la mécanique quantique. On se servira de la terminologie des physiciens pour décrire les propriétés des solutions de l'équation sans s'arrêter sur les motivations physiques des définitions introduites.

A. Définitions et notations.

En physique, l'équation

$$\frac{d^2\Psi}{dx^2} + (E - U(x))\Psi = 0 \tag{1}$$

est appelée *équation stationnaire de Schrödinger*.

La variable indépendante x est la *coordonnée cartésienne d'une particule*. La fonction inconnue Ψ qui est généralement complexe est appelée *fonction d'onde de la particule*; les solutions de l'équation de Schrödinger, *états de la particule*; le paramètre spectral E, *énergie de la particule*; la fonction connue U, *potentiel* ou *énergie potentielle de la particule*. La *mécanique quantique* étudie essentiellement les propriétés de l'équation (1) et les équations et systèmes d'équations à dérivées partielles la généralisant.

E x e m p l e. Soit $U = 0$. La particule est dite *libre*. L'equation de Schrödinger pour une particule libre d'énergie $E = k^2$ s'écrit

$$\Psi_{xx} + k^2\Psi = 0. \tag{2}$$

Cette équation possède deux solutions linéairement indépendantes

$$\Psi_+ = e^{ikx}, \ \Psi_- = e^{-ikx}.$$

Ces deux solutions sont dites *particule se déplaçant à droite* (avec une impulsion $k > 0$) et respectivement *particule se déplaçant à gauche*. Donc, l'espace des états d'une particule libre d'énergie E est un espace complexe à deux dimensions.

Le carré du module de la fonction d'onde est appelé *densité de probabilité* de trouver la particule en un endroit donné. Donc, une particule libre d'impulsion k se trouve « avec la même probabilité en un point quelconque » (on peut utiliser cette terminologie sans se soucier de la signification de ces termes et comment ils sont rattachés à la théorie des probabilités).

B. Barrières potentielles.

Admettons que le potentiel est à support borné (non nul seulement dans un domaine borné). On dit qu'est donnée une *barrière*

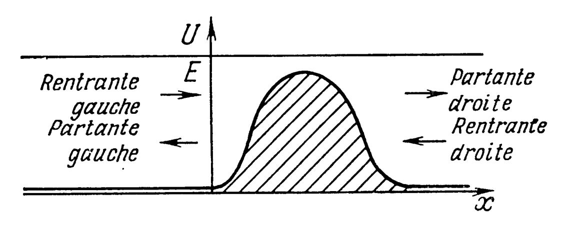


Fig. 29

potentielle si $U \geqslant 0$, et une *fosse potentielle* si $U \leqslant 0$. Le domaine où le potentiel est non nul s'appelle *support* du potentiel (fig. 29).

Supposons que l'énergie de la particule est *positive*: $E = k^2$. L'équation de Schrödinger (1) est alors confondue à gauche du support avec l'équation de la particule libre (2). Donc elle possède deux solutions qui sont confondues *à gauche du support* respectivement avec e^{ikx} et e^{-ikx}. Ces deux solutions sont dites *particules rentrante gauche* et *partante gauche* respectivement. On remarquera que ces solutions sont définies pour tous les x, mais ne sont confondues avec e^{ikx} et e^{-ikx} qu'à gauche du support.

De façon analogue, il existe deux solutions confondues avec e^{ikx} et e^{-ikx} *à droite du support*. Ces solutions sont appelées *particules partante droite* et *rentrante droite* respectivement.

E x e r c i c e. Une particule rentrante gauche peut-elle être entièrement réfléchie à gauche (c'est-à-dire une fonction d'onde peut-elle être nulle à droite de la barrière et non nulle à gauche)? Passer entièrement à droite?

R é p o n s e. Non, oui.

C. Opérateur de monodromie.

D é f i n i t i o n. On appelle *opérateur de monodromie* de l'équation de Schrödinger (1) à potentiel à support borné un opérateur linéaire de l'espace des états d'une particulle *libre* d'énergie $E = k^2$ dans lui-même défini comme suit.

A chaque solution de l'équation de la particule libre (2) est associée la solution de l'équation de Schrödinger qui est confondue avec elle à gauche du support, et à cette solution, sa valeur à droite du support.

On constate que l'opérateur de monodromie est doué de remarquables propriétés de (1,1)-unitarité. Pour les énoncer on aura besoin des

N o t a t i o n s. Désignons par $\mathbb{R}^2$ l'*espace des solutions réelles de l'équation de Schrödinger* (1). *L'espace des états de la particule* (i.e. l'espace des solutions complexes) est le complexifié de $\mathbb{R}^2$. On le notera $\mathbb{C}^2 = {}^{\mathbb{C}}\mathbb{R}^2$. Cet espace renferme les quatre états de la particule rentrante et partante gauche et droite.

On désignera par $\mathbb{R}_0^2$ *l'espace des solutions réelles de l'équation de la particule libre* (2) (puisque $U = 0$ pour une particule libre). Cet espace est muni de la base naturelle

$$e_1 = \cos kx, \quad e_2 = \sin kx.$$

On désignera par $\mathbb{C}_0^2$ l'*espace des états de la particule libre*; c'est le complexifié de l'espace $\mathbb{R}_0^2$. Une base naturelle est constituée des états des particules se déplaçant à droite et à gauche respectivement. On les notera

$$f_1 = e^{ikx}, \quad f_2 = e^{-ikx}.$$

A remarquer que e_1 et e_2 forment aussi une base dans l'espace des états. Entre ces deux bases on a la relation

$$f_1 = e_1 + ie_2, \quad f_2 = e_1 - ie_2.$$

D é f i n i t i o n. Le groupe SU (1, 1) des *matrices* (1, 1)-*unitaires unimodulaires* est composé de toutes les matrices d'ordre deux, de déterminant un, préservant la forme hermitienne $|z_1|^2 - |z_2|^2$. En d'autres termes, ce sont des matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ telles que

$$|a|^2 - |b|^2 = |c|^2 - |d|^2 = 1, \quad a\bar{c} - b\bar{d} = 0, \quad ad - bc = 1.$$

T h é o r è m e. *La matrice d'un opérateur de monodromie dans la base* (f_1, f_2) *appartient au groupe* SU (1, 1).

L'appartenance d'un opérateur de monodromie au groupe SU (1, 1) est due au fait que le flot de l'équation (1) préserve les aires. Pour le prouver nous allons rappeler quelques généralités sur le groupe SU (1, 1).

D. Digression algébrique: le groupe SU(1, 1).

Soient l'espace vectoriel réel $\mathbb{R}^2$ et son complexifié $\mathbb{C}^2$. Considérons l'élément de surface de $\mathbb{R}^2$ et désignons par $[\xi, \eta]$ l'aire orientée du parallélogramme construit sur les vecteurs ξ et η. On appellera *structure symplectique* le produit scalaire gauche [,]. Si dans $\mathbb{R}^2$ est fixée une base (e_1, e_2) dans laquelle $[e_1, e_2] = 1$, alors $[\xi, \eta]$ est égal au déterminant composé des coordonnées des vecteurs ξ et η dans la base (e_1, e_2).

La complexifiée de la forme bilinéaire [,] définit une structure symplectique dans $\mathbb{C}^2$; nous la désignerons par les mêmes crochets.

A noter que la forme [,] est régulière: si $[\xi, \eta] = 0$ pour tous les ξ, alors $\eta = 0$.

Soit dans $\mathbb{C}^2$ la *forme hermitienne* $\langle \xi, \eta \rangle = \frac{i}{2} [\xi, \bar{\eta}]$. Cette forme est effectivement hermitienne: $\langle \lambda\xi, \eta \rangle = \lambda \langle \xi, \eta \rangle$, $\langle \xi, \eta \rangle = \overline{\langle \eta, \xi \rangle}$. Dans la suite, on a intérêt à calculer les produits hermitiens des vecteurs $f_1 = e_1 + ie_2$ et $f_2 = e_1 - ie_2$. Il est immédiat de vérifier le

L e m m e. *On a les relations*

$$\langle f_1, f_1 \rangle = 1, \quad \langle f_2, f_2 \rangle = -1, \quad \langle f_1, f_2 \rangle = 0.$$

◀ Par exemple,

$$\langle f_1, f_1 \rangle = \frac{i}{2} [f_1, \bar{f}_1] = \frac{i}{2} [f_1, f_2] = \frac{i}{2} \begin{vmatrix} 1 & i \\ 1 & -i \end{vmatrix} = 1. \ ▶$$

Donc, la forme hermitienne ⟨,⟩ est du « type (1, 1) » (un carré positif et un carré négatif dans la forme canonique $\langle z, z \rangle = |z_1|^2 - |z_2|^2$).

Considérons maintenant les transformations linéaires du plan $\mathbb{C}^2$ préservant les structures hermitienne, symplectique et réelle.

D é f i n i t i o n. On appellera *groupe* (1, 1)*-unitaire* et on le notera $U(1, 1)$ le groupe des applications linéaires du plan $\mathbb{C}^2$.

Le groupe des applications linéaires du plan $\mathbb{C}^2$, préservant la structure symplectique [,] est appelé *groupe linéaire spécial* (*ou unimodulaire*) *d'ordre deux* et noté SL (2, $\mathbb{C}$).

Le groupe des applications linéaires réelles du plan $\mathbb{C}^2$ (c'est-à-dire le groupe dont les éléments sont les complexifiées des applications linéaires de $\mathbb{R}^2$) s'appelle *groupe linéaire réel du second ordre* et se note GL (2, $\mathbb{R}$).

Nous venons donc de définir trois sous-groupes du groupe GL (2, $\mathbb{C}$) des applications linéaires de $\mathbb{C}^2$: le groupe (1, 1)-unitaire U (1, 1), le groupe unimodulaire SL (2, $\mathbb{C}$) et le groupe réel GL (2, $\mathbb{R}$).

En outre, entre la forme hermitienne ⟨,⟩ définissant le groupe unitaire, la forme symplectique [,] définissant le groupe unimodulaire et la conjugaison complexe définissant le groupe réel, on a la rela-

tion

$$\langle a, b\rangle = \frac{i}{2}[a, \bar{b}].$$

T h é o r è m e. *L'intersection de deux quelconques de ces trois sous-groupes est confondue avec l'intersection de tous les trois* (fig. 30).

Cette intersection s'appelle *groupe* (1, 1)*-unitaire spécial* *) et se note SU (1, 1). (On le nomme également groupe réel unimodulaire et on le note SL (2, $\mathbb{R}$). On l'appelle encore groupe symplectique du second ordre et on le désigne par Sp (1, $\mathbb{R}$).)

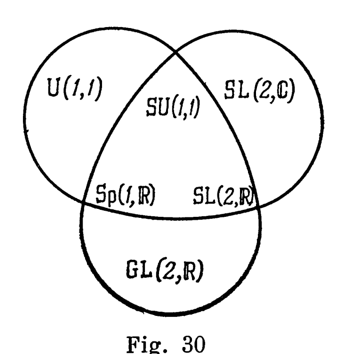


Fig. 30

◀ Si une application A est réelle et unimodulaire, alors

$$[A\xi, A\eta] = [\xi, \eta] \text{ et } A\bar{\xi} = \overline{A\xi}.$$

Donc

$$\langle A\xi, A\eta\rangle = \frac{i}{2}[A\xi, \overline{A\eta}] = \frac{i}{2}[A\xi, A\bar{\eta}] = \frac{i}{2}[\xi, \eta] = \langle \xi, \eta\rangle.$$

Si A est réelle et (1, 1)-unitaire, alors

$$\langle A\xi, A\eta\rangle = \langle \xi, \eta\rangle \text{ et } A\bar{\xi} = \overline{A\xi}.$$

Donc

$$[A\xi, A\eta] = -2i\langle -A\xi, \overline{A\eta}\rangle = -2i\langle A\xi, A\bar{\eta}\rangle = -2i\langle \xi, \bar{\eta}\rangle = [\xi, \eta].$$

Si A est (1, 1)-unitaire et unimodulaire, alors

$$[A\xi, A\eta] = [\xi, \eta] \text{ et } [A\xi, \overline{A\eta}] = [\xi, \bar{\eta}],$$

donc

$$[A\xi, \overline{A\eta}] = [A\xi, A\bar{\eta}], \quad \forall\xi, \forall\eta.$$

*) On rappelle qu'il s'agit d'un groupe d'opérateurs et non de matrices. Les matrices de ces opérateurs appartiennent au groupe de matrices SU (1, 1) dans la base spécialement choisie plus haut.

Par suite,

$$[\xi, \overline{A\eta} - A\bar{\eta}] = 0, \quad \forall \xi$$

et

$$\overline{A\eta} = A\bar{\eta}, \quad \forall \eta,$$

c'est-à-dire A est réelle. ▶

Corollaire. *Si la matrice d'un opérateur est réelle et unimodulaire dans la base réelle* (e_1, e_2), *elle est spéciale* (1, 1)*-unitaire dans la base conjuguée complexe* $(f_1 = e_1 + ie_2,\ f_2 = e_1 - ie_2)$ *et inversement.*

◀ Le carré scalaire hermitien du vecteur $z_1 f_1 + z_2 f_2$ a pour expression

$$\langle z, z\rangle = |z_1|^2 - |z_2|^2$$

où (z_1, z_2) sont les coordonnées de ce vecteur dans la base (f_1, f_2) (voir lemme). On a donc
(matrice de A réelle et unimodulaire dans la base (e_1, e_2)) $\Leftrightarrow$
$\Leftrightarrow$ $(A \in GL(2, \mathbb{R}) \cap SL(2, \mathbb{C}))$ $\Leftrightarrow$ $(A \in SU(1, 1))$ $\Leftrightarrow$ (matrice de A (1, 1)-unitaire et unimodulaire dans la base (f_1, f_2)). ▶

E. Digression géométrique : SU(1, 1) et géométrie de Lobatchevsky.

Les groupes de matrices SL (2, $\mathbb{R}$) et SU (1, 1) sont rattachés de la manière suivante à la géométrie de Lobatchevsky (fig. 31).

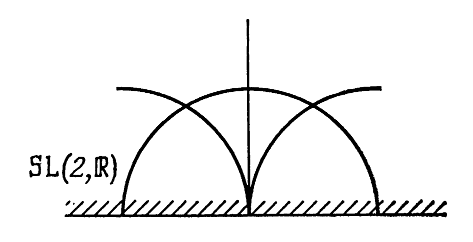


Fig. 31

Une matrice unimodulaire réelle d'ordre deux définit une application homographique $z \mapsto (az + b)/(cz + d)$ du demi-plan supérieur dans lui-même. Cette application est un mouvement du plan de Lobatchevsky réalisé comme le demi-plan supérieur. Tous les mouvements du plan de Lobatchevsky s'obtiennent de cette manière. Le groupe des mouvements du plan de Lobatchevsky est isomorphe à SL (2, $\mathbb{R}$)/$\pm E$.

Une matrice (1, 1)-unitaire unimodulaire du second ordre définit une application homographique du disque unité dans lui-

même. En effet, l'application (1, 1)-unitaire unimodulaire laisse invariant le cône $|z_1|^2 < |z_2|^2$. L'application naturelle

$$\mathbb{C}^2 \setminus O \to \mathbb{C}P^1, \quad (z_1, z_2) \mapsto w = z_1 / z_2$$

envoie ce cône dans le disque unité $|w| < 1$ et les applications linéaires de $\mathbb{C}^2$ dans les applications homographiques de $\mathbb{C}P^1$ (fig. 32).

Les applications homographiques du disque unité dans lui-même définies par les matrices de SU (1, 1) sont des mouvements du plan de Lobatchevsky réalisé comme l'intérieur du disque unité. Tous les mouvements du plan de Lobatchevsky s'obtiennent de la sorte. Le groupe des mouvements du plan de Lobatchevsky est isomorphe au groupe SU $(1, 1)/\pm E$.

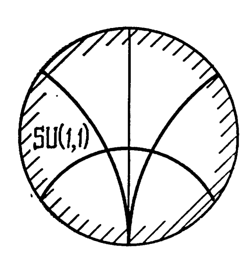


Fig. 32

Les groupes de matrices SL $(2, \mathbb{R})$ et SU $(1, 1)$ sont isomorphes: ils se déduisent à partir d'un même groupe d'opérateurs. Les matrices des opérateurs exprimées dans la base réelle (e_1, e_2) appartiennent à SL $(2, \mathbb{R})$ et dans la base conjuguée complexe (f_1, f_2) au groupe SU $(1, 1)$. Le passage de SL $(2, \mathbb{R})$ à SU $(1, 1)$ correspond au passage de la base réelle à la base conjuguée complexe et du modèle du plan de Lobatchevsky sur le demi-plan supérieur au modèle sur le disque unité.

Exercice. Montrer que le groupe SL $(2, \mathbb{R})$ est homéomorphe à $S^1 \times D^2$ (l'intérieur d'un tore).

F. Propriétés de l'opérateur de monodromie réelle.

Revenons à l'opérateur de monodromie de l'équation de Schrödinger (1). Outre les espaces $\mathbb{R}^2$ et $\mathbb{R}^2_0$ des solutions de l'équation (1) et de l'équation d'une particule libre nous allons considérer le *plan de phase* $\mathbb{R}^2_\Phi$ dont les points sont des couples de nombres réels (Ψ, Ψ_x).

Soit $x \in \mathbb{R}$. Considérons un opérateur linéaire associant à toute solution (réelle) de l'équation (1) sa condition initiale au point x

$$B^x \colon \mathbb{R}^2 \mapsto \mathbb{R}^2_\Phi, \quad \Psi \mapsto (\Psi(x), \Psi_x(x)).$$

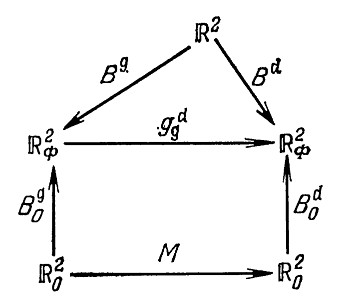
Cet opérateur définit un isomorphisme. L'isomorphisme $g_{x_1}^{x_2} = B^{x_2}(B^{x_1})^{-1}$ s'appelle *application de phase de* x_1 à x_2.

L'opérateur (solution $\mapsto$ point de phase)

$$B_0^x \colon \mathbb{R}^2_0 \to \mathbb{R}^2_\Phi$$

se définit de façon analogue pour l'équation de la particule libre (2).

Dans ces notations l'opérateur réel de monodromie M est défini par le diagramme commutatif d'isomorphismes:



g désigne le point x à gauche du support, d, à droite. L'opérateur M est indépendant du choix de ces points.

Théorème. *Le déterminant de l'opérateur de monodromie de l'équation de Schrödinger est égal à un.*

◀ L'espace $\mathbb{R}^2_0$ des états réels de la particule libre est muni de la base $e_1 = \cos kx$, $e_2 = \sin kx$. L'espace des phases réel $\mathbb{R}^2_\Phi$ est rapporté à un système de coordonnées Ψ, Ψ_x. Donc il est muni d'une base dans laquelle la matrice de l'opérateur s'écrit

$$(B_0^x) = \begin{pmatrix} \cos kx & \sin kx \\ -k\sin kx & k\cos kx \end{pmatrix}.$$

Par suite, $\det B_0^x = k$ ne dépend pas de x. En particulier, les déterminants des isomorphismes verticaux de gauche et de droite du diagramme sont égaux. Donc, $\det M = \det g_g^d$ (le diagramme est commutatif). Or le flot conserve les aires d'après le théorème de Liouville (l'équation de Schrödinger ne contient pas de terme en Ψ_x). Donc, $\det g_g^d = 1$. Et, par suite, $\det M = 1$. ▶

G. Propriétés de l'opérateur de monodromie complexe.

◀ Démonstration du théoreme de (1, 1)-unitarité du point C. L'opérateur complexe de monodromie est le complexifié de l'opérateur réel (cf. point C).

La matrice de l'opérateur de monodromie dans la base réelle (e_1, e_2) appartient à SL (2, $\mathbb{R}$) (voir point F). Donc dans la base conjuguée complexe ($f_1 = e_1 + ie_2$, $f_2 = e_1 - ie_2$) la matrice de cet opérateur appartient à SU (1, 1) (cf. point D, corollaire). ▶

Exercice. Montrer que l'équation de Schrödinger (1) ne peut posséder de solution non nulle confondue avec ae^{ikx} à gauche du support et avec be^{-ikx} à droite du support (la particule ne peut rentrer sans partir).

Solution. L'opérateur de monodromie conserve le carré (1, 1)-hermitien $|z_1|^2 - |z_2|^2$. Or $\langle ae^{ikx}, ae^{ikx}\rangle = |a|^2$, $\langle be^{-ikx}, be^{-ikx}\rangle = -|b|^2$ (cf. lemme du point D). Par suite, $|a|^2 = -|b|^2$, i.e. $a = b = 0$.

H. Coefficients de passage et de réflexion.

D é f i n i t i o n. On dit qu'une particule rentrante gauche avec une impulsion $k > 0$ traverse une barrière avec un *coefficient de passage* $|A|^2$ et un *coefficient de réflexion* $|B|^2$ si l'équation de Schrödinger (1) avec $E = k^2$ admet une solution Ψ égale à

$$e^{ikx} + Be^{-ikx} \text{ à gauche de la barrière,}$$

$$Ae^{ikx} \text{ à droite de la barrière (fig. 33).}$$

L e m m e. *La solution Ψ et les constantes complexes A et B réalisant les conditions décrites existent et sont uniques quel que soit $k > 0$.*

◀ Soit une particule partante droite (la solution e^{ikx} à droite de la barrière). A gauche de la barrière cette solution de même que toute autre solution se présente sous la forme d'une combinaison linéaire de e^{ikx} et e^{-ikx}, le coefficient en e^{ikx} étant non nul par suite de la (1, 1)-unitarité de l'opérateur de monodromie (cf. problème du point G).

Une division par ce coefficient non nul nous donne la solution cherchée. Donc les coefficients A et B sont définis de façon univoque. ▶

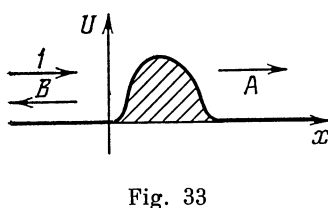


Fig. 33

E x e r c i c e. Montrer que le coefficient de passage n'est jamais nul.
S o l u t i o n. Si $\Psi \equiv 0$ à droite de la barrière, elle l'est à gauche.

T h é o r è m e. *La somme des coefficients de passage et de réflexion est égale à un.*

◀ L e m m e. *Dans la base ($f_1 = e^{ikx}$, $f_2 = e^{-ikx}$) la matrice de l'opérateur de monodromie s'exprime en fonction des coefficients complexes A et B par la formule*

$$(M) = \begin{pmatrix} 1/\overline{A} & -\overline{B}/\overline{A} \\ -B/A & 1/A \end{pmatrix}.$$

◀ Par définition des coefficients A et B, l'opérateur de monodromie opère de telle sorte que $f_1 + Bf_2 \mapsto Af_1$. Cet opérateur étant réel nous pouvons trouver l'image de son conjugué complexe. Comme $\overline{f_1} = f_2$, il vient $f_2 + \overline{B}f_1 \mapsto \overline{A}f_2$. En divisant par $A \neq 0$ et $\overline{A} \neq 0$ on obtient la matrice de l'inverse de l'opérateur de monodromie

$$(M^{-1}) = \begin{pmatrix} 1/A & \overline{B}/\overline{A} \\ B/A & 1/\overline{A} \end{pmatrix}.$$

[Pour inverser une matrice unimodulaire du second ordre il suffit de permuter les éléments diagonaux et de changer le signe

des autres:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}=\begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix}.] \blacktriangleright$$

Les conditions $M \in SU(1, 1)$ donnent $1/|A|^2 - |B|^2/|A|^2 = 1$. ▶

E x e r c i c e. Calculer les coefficients de passage et de réflexion pour unipotentiel égal à une constante U_0 pour $0 \leqslant x \leqslant a$ et à zéro ailleurs (fig. 34).

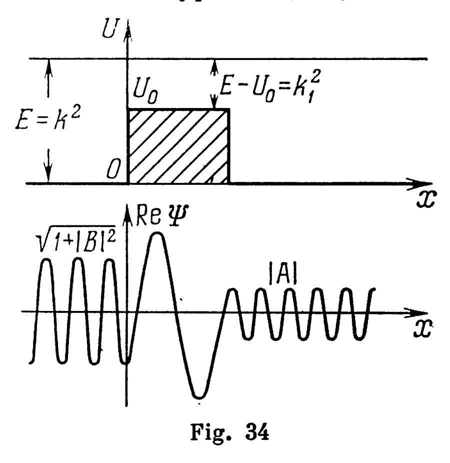


Fig. 34

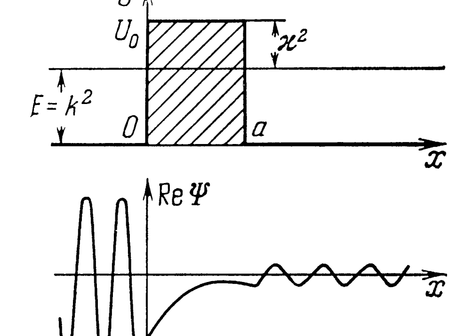


Fig. 35

R é p o n s e.

$$|A|^2=\frac{1}{1+\frac{U_0^2 \sin^2 ak_1}{4E(E-U_0)}},$$

où $E = k^2$, $E - U_0 = k_1^2$ (la particule est freiné en franchissant la barrière, donc la densité de probabilité est plus grande de la trouver qu'en dehors de la

barrière). Pour E grand, le coefficient de réflexion tend vers 0

$$|B|^2 \sim \frac{U_0^2}{4E\,(E-U_0)} \sin^2 ak_1.$$

Si l'énergie de la particule est numériquement inférieure à la hauteur de la barrière, le coefficient de passage est exponentiellement petit:

$$|A|^2 = \frac{4k^2\varkappa^2}{(k^2+\varkappa^2)\,\mathrm{sh}^2\, a\varkappa + 4k^2\varkappa^2},$$

où $E = k^2$, $U_0 - E = \varkappa^2$ (fig. 35). Le coefficient de passage au-dessus d'une barrière large et haute, même s'il est petit, est toujours distinct de zéro (« effet de tunnel »): la particule quantique « passe sous la barrière » qui est infranchissable par une particule classique).

I. Matrice de dispersion.

On étudie également la traversée de la barrière de droite à gauche. La solution correspondante Ψ_2 est égale à

$$e^{-ikx} + B_2 e^{ikx} \text{ à droite de la barrière,}$$
$$A_2 e^{-ikx} \text{ à gauche (fig. 36).}$$

D é f i n i t i o n. On appelle *matrice de dispersion* (ou S-matrice) a matrice

$$S = \begin{pmatrix} A & B \\ B_2 & A_2 \end{pmatrix}.$$

R e m a r q u e. A partir de l'équation stationnaire de Schrödinger il est difficile de comprendre quel opérateur est associé à cette matrice et pourquoi cette matrice est douée des propriétés remarquables que nous allons démontrer. Ceci s'explique par le fait que S « transforme les particules rentrantes en particules partantes ». Cette phrase prend toute sa signification si l'on considère l'équation non stationnaire (ce que nous ne ferons pas).

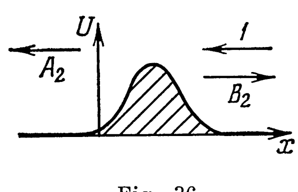


Fig. 36

T h é o r è m e. *La matrice de dispersion est unitaire et, de plus, les coefficients de passage de gauche à droite et de droite à gauche sont égaux*: $A = A_2$.

◀ L'opérateur de monodromie opère comme

$$A_2 f_2 \mapsto f_2 + B_2 f_1, \quad \overline{A}_2 f_1 \mapsto f_1 + \overline{B}_2 f_2.$$

Donc,

$$(M) = \begin{pmatrix} 1/\overline{A}_2 & B^2/A_2 \\ \overline{B}_2/\overline{A}_2 & 1/A_2 \end{pmatrix}.$$

Une comparaison avec la matrice calculée dans le lemme du point H donne

$$A_2 = A, \quad B_2 = -\bar{B}A/\bar{A}.$$

Comme $|A|^2 + |B|^2 = |A_2|^2 + |B_2|^2 = 1$ et $A\bar{B}_2 + B\bar{A}_2 = 0$, la matrice S est unitaire. ▶

R e m a r q u e. Nous avons étudié l'équation de Schrödinger (1) pour les valeurs réelles du paramètre spectral $E = k^2$. L'étude des valeurs complexes de k est très utile. Il se trouve que les propriétés d'unitarité et de symétrie de la matrice de dispersion ont lieu pour k complexe. De plus, S jouit de la propriété d'« être réel », $S(-k) = \overline{S(k)}$ et d'« analyticité »: $A(k)$ est la valeur limite d'une fonction analytique dans le demi-plan imaginaire $\text{Im}\, k > 0$, présentant un nombre fini de pôles sur l'axe imaginaire.

Les coefficients de passage et de réflexion étant mesurables, on peut poser le problème inverse de la théorie de dispersion qui consiste à déterminer le potentiel U sachant la matrice de dispersion $S(k)$.

Le potentiel U est défini par une fonction réelle sur la droite réelle ou par deux fonctions réelles sur la demi-droite. Les coefficients A et B sont deux fonctions complexes sur le demi-axe $k > 0$, c'est-à-dire quatre fonctions réelles sur le demi-axe. La condition d'unitarité $|A|^2 + |B|^2 = 1$ réduit de quatre à trois le nombre de fonctions réelles sur le demi-axe.

Comme $3 > 2$, tout couple (A, B) tel que $|A|^2 + |B|^2 = 1$ n'est pas forcément associé à un potentiel; pour déterminer le potentiel sachant A et B, il faut imposer à ces coefficients encore une condition qui n'est autre que la condition d'analyticité.

Le plus étonnant c'est que ces généralités euristiques grossières, basées sur le calcul du nombre de fonctions arbitraires ont été traduites par des théorèmes exactement formulés et démontrés. Ces derniers sortent du cadre de notre cours.

J. Etats liés.

Soit maintenant un potentiel ayant la forme d'une fosse à support borné ($U(x) \leqslant 0$, $U(\infty) = 0$). On dit qu'une particule *se trouve dans la fosse* si $\Psi \to 0$ pour $x \to \pm\infty$ (fig. 37). Il est clair qu'une particule ne peut se trouver dans la fosse que si son énergie E est *négative*. A gauche et à droite de la fosse la solution est une combinaison linéaire des exponentielles $e^{\varkappa x}$ et $e^{-\varkappa x}$, où $\varkappa^2 = -E$, $\varkappa > 0$. Par suite, une condition pour que la particule se trouve dans la fosse est que le coefficient en l'exponentielle croissante à gauche s'annule à gauche de la fosse et celui de l'exponentielle croissante à droite, à droite de la fosse. Les solutions douées de telles propriétés n'existent pas pour toute valeur négative de l'énergie.

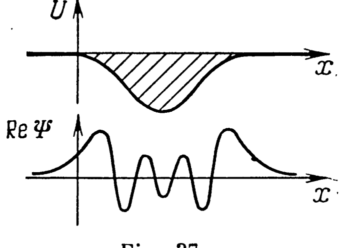


Fig. 37

Si la fosse est assez profonde et large, il existe un nombre fini de valeurs négatives de l'énergie E telles que la particule se trouve

à l'état stationnaire dans la fosse. Ces valeurs sont d'autant plus nombreuses que la fosse est profonde et large.

Les valeurs correspondantes de E s'appellent *niveaux stationnaires*, les fonctions d'onde Ψ *amorties pour* $x \to \pm \infty$, *états liés* (si la fosse n'est pas à support borné, on exige que $\int |\Psi|^2 dx < \infty$).

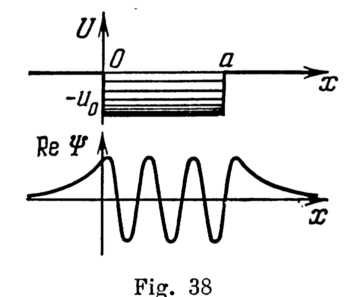


Fig. 38

E x e r c i c e. Déterminer les niveaux stationnaires de l'énergie dans une fosse rectangulaire d'une profondeur U_0, comprise entre $x = 0$ et $x = a$ (fig. 38).

R é p o n s e. $E = 4\xi^2/a^2 - U_0$, où ξ sont racines des équations

$$\cos \xi = \pm \gamma\xi, \quad \sin \xi = \pm \gamma\xi \quad \left(\gamma = \frac{1}{a} \sqrt{4/U_0}\right)$$

($\operatorname{tg} \xi > 0$ pour la première équation, $\operatorname{tg} \xi < 0$ pour la seconde).

E x e r c i c e. Prouver que les fonctions d'onde Ψ_j correspondant à des états liés ayant des niveaux d'énergie distincts sont orthogonales: $\int \Psi_1 \overline{\Psi}_2 \, dx = 0$.

E x e r c i c e *. Etablir une relation entre les niveaux stationnaires de l'énergie d'états liés et les pôles d'une S-matrice sur l'axe imaginaire k.

6. Géométrie de l'équation différentielle du second ordre et géométrie d'un couple de champs de directions dans un espace à trois dimensions

On étudie les propriétés locales des solutions de l'équation différentielle du second ordre qui sont géométriques, c'est-à-dire invariantes par les difféomorphismes du plan des variables dépendante et indépendante.

Chaque équation différentielle du second ordre est liée à un couple de champs de directions dans un espace à trois dimensions. Le problème de la classification locale des équations du second ordre aux difféomorphismes près du plan est équivalent à celui de la classification locale d'un couple de champs de directions génériques dans l'espace aux difféomorphismes près de cet espace. Plus bas on étudie les invariants et les « formes normales » de ces deux problèmes équivalents.

A. Propriétés configuratives des solutions des équations linéaires. Les graphes des solutions de l'équation $d^2y/dx^2 = 0$ (des droites) vérifient les théorèmes de configuration (de Papp, de Désargues, etc.) de la géométrie projective.

T h é o r è m e. *La famille des graphes des solutions de toute équation homogène linéaire du second ordre*

$$d^2y/dx^2 + a(x)\, dy/dx + b(x)\, y = 0$$

*est localement (au voisinage de tout point $x = x_0$) difféomorphe *) à la famille des graphes des solutions de l'équation simple $d^2y/dx^2 = 0$.*

---

*) i.e. il existe un difféomorphisme du voisinage de la droite $x = x_0$ sur le plan (x, y) envoyant les graphes des solutions dans des droites.

C o r o l l a i r e. *Les théorèmes de configuration de la géométrie projective sont valables (localement) pour les graphes des solutions de toute équation linéaire du second ordre, par exemple pour la famille de courbes* $y = A \sin x + B \cos x$ *ou* $y = Ae^x + Be^{-x}$.

◀ Considérons une solution y_1 non nulle en x_0 et une autre solution y_2 nulle en x_0 mais non identiquement nulle. Les formules

$$Y = y/y_1, \quad X = y_2/y_1$$

définissent le difféomorphisme cherché. ▶

R e m a r q u e 1. Les coordonnées (X, Y) sont définies à une transformation homographique près (le changement des solutions y_1 et y_2 par leurs combinaisons linéaires). En particulier, la coordonnée X définit sur l'axe x une structure de variété localement projective d'une variété à une dimension (un atlas dans lequel les fonctions de raccordement sont des applications projectives de la droite, i.e. des fonctions homographiques).

De façon analogue, l'équation homogène linéaire du second ordre définit sur le plan (x, y) une structure de plan localement projectif.

R e m a r q u e 2. On appelle *équivalence* de variétés localement projectives un difféomorphisme transformant une structure localement projective dans une autre.

E x e r c i c e. Enumérer les diverses structures d'une variété localement projective à l'équivalence près:

a) sur la droite;

b) sur le cercle.

I n d i c a t i o n. Les structures localement projectives de la droite sont induites par une application dans une droite projective (i.e. dans un cercle) avec une dérivée non nulle; le nombre des antécédents d'un point du cercle par cette application est un invariant de la structure.

Le revêtement à deux feuillets de la droite projective définit sur le cercle une structure de variété localement projective non équivalente à la structure de la droite projective. Cependant toute structure localement projective sur le cercle n'est pas induite à partir de la structure de la droite projective. La classification des structures localement projectives du cercle est liée à celle des équations de Hill (équations linéaires du second ordre à coefficients périodiques). Les équations à coefficients constants définissent déjà des structures non induites à droite projective.

B. Forme normale de la partie quadratique de l'équation du second ordre au voisinage d'une solution donnée.

Soit maintenant une équation non linéaire du second ordre

$$d^2y/dx^2 = \Phi(x, y, dy/dx).$$

On se propose d'étudier la géométrie de la famille à deux paramètres de courbes, définie par cette équation sur le plan (x, y). On s'intéresse plus particulièrement à la validité des théorèmes de configuration dans cette famille et à la possibilité de rectifier cette famille (c'est-à-dire de la transformer en une famille de droites) par un difféomorphisme convenable du plan. On verra que la rectification n'est pas toujours possible et on indiquera des invariants mesurant le degré de violation du théorème de Désargues.

T h é o r è m e. *Au voisinage de chaque élément linéaire (x, y, p) du plan des variables dépendante et indépendante, toute équation différentielle du second ordre est susceptible d'être ramenée par un difféomorphisme local de ce plan à la forme*

$$d^2y/dx^2 = A(x)\, y^2 + O(|y|^3 + |p|^3); \quad p = dy/dx$$

au voisinage de l'élément $(x = 0,\ y = 0,\ p = 0)$.

◀ 1°. *Elimination des termes linéaires.*

L'élément linéaire donné définit l'unique solution dont le graphe peut être pris pour axe x. Linéarisons cette équation au voisinage de cette solution. L équation linéaire obtenue peut être en vertu du théorème du point A, localement rectifiée (ramenée à la forme $d^2y/dx^2 = 0$) par un choix *) convenable du système de coordonnées. Dans ce système, l'équation du second ordre a pour second membre $\Phi(x, y, p)$ qui s'annule pour $y = 0$, $p = 0$ avec les dérivées par rapport à y et p. Dans ce système, le développement de Φ en série de Taylor sur y et p commence par des termes de degré non inférieur à deux

$$d^2y/dx^2 = A(x)\,y^2 + B(x)\,yp + C(x)\,p^2 + O(|y|^3 + |p|^3).$$

2°. *Transformation de la variable indépendante.*

Soit un difféomorphisme local du plan (x, y), défini par le changement

$$x = F(X, Y),\ y = Y$$

(qui associe à un point de coordonnées (x, y) le point de coordonnées (X, Y)).

L e m m e. L'équation

$$d^2y/dx^2 = \Phi(x, y, p), \qquad p = dy/dx,$$

se transforme par ce changement en l'équation

$$d^2Y/dX^2 = \hat{\Phi}(X, Y, P),\ P = dY/dX,$$

où

$$\hat{\Phi}(X, Y, P) = \frac{\Delta^3}{F'}\Phi(F, Y, P/\Delta) + \frac{F'' + 2PF'_Y + P^2F_{YY}}{F'}P;$$

$\Delta = F' + PF_Y$, *le symbole prime représente la dérivée partielle par rapport à* X; F *et ses dérivées sont fonctions de* X *et* Y.

◀ Soient $y = u(x)$ solution de l'équation initiale, $Y = U(X)$ son image. On a

$$P = \frac{dU}{dX} = \frac{du}{dx}\Big|_F \cdot \left(F' + PF_Y\right),\ \frac{du}{dx}\Big|_{F(X,\,U,\,(X))} = \frac{P}{\Delta}.$$

Donc,

$$\frac{d^2U}{dX^2} = \Delta\,\frac{d}{dX}\,\frac{du}{dx}\Big|_{F(X,\,U(X))} + \frac{P}{\Delta}\left(F'' + 2F'_YP + F_{YY}P^2 + F_Y\frac{d^2U}{dX^2}\right).$$

Ensuite

$$\frac{d}{dX}\,\frac{du}{dx}\Big|_{F(X,\,U(X))} = \frac{d^2u}{dx^2}\Big|_F\,\Delta = \Phi(F(X, U(X)), U(X), P/\Delta)\,\Delta.$$

En définitive,

$$\left(1 - \frac{PF_Y}{\Delta}\right)\frac{d^2U}{dX^2} = \Phi\,(F, U, P/\Delta)\,\Delta^2 + \frac{P}{\Delta}(F'' + 2PF'_Y + P^2F_{YY}),$$

d'où l'on déduit la formule annoncée. ▶

Cette formule entraîne aussitôt le

C o r o l l a i r e 1. *Soit* $\Phi = 0$. *Alors* $\hat{\Phi}$ *est un polynôme en* p *de degré* $\leqslant 3$.

C o r o l l a i r e 2. *Soit* Φ *un polynôme en* p *de degré* $\leqslant 3$. *Alors* $\hat{\Phi}$ *est un polynôme en* P *de degré* $\leqslant 3$.

R e m a r q u e. Les images des polynômes en p de degré n, $n \geqslant 4$, par l'application $\Phi \mapsto \hat{\Phi}$ ne sont plus des polynômes en P. En effet, $\Delta^3(P/\Delta)^n$ n'est pas un polynôme en P pour $n > 3$.

*) Pour la suite, il est utile de remarquer que si le second membre de Φ est un polynôme en p de degré $n \geqslant 1$, cette propriété reste valable dans les nouvelles coordonnées construites au point A.

C o r o l l a i r e 3. *L'équation différentielle du second ordre définit sur le graphe de chacune de ses solutions une structure de droite localement projective et sur le fibré normal au graphe une structure de plan localement projectif.*

◀ Etudions l'équation aux variations le long d'une solution donnée. C'est une équation homogène linéaire du second ordre; donc sur le plan de la variable dépendante et de la variable indépendante est définie une structure de plan localement projectif et sur la solution une structure de droite localement projective (cf. point A).

Dans les notations du lemme l'équation aux variations s'écrit $d^2y/dx^2 = \Phi_1$ où $\Phi = \Phi_1 + \Phi_2 + \ldots$ est le développement en série sur y et p.

Considérons maintenant le fibré normal du graphe de la solution étudiée sur le plan. L'espace normal à une sous-variété en l'un quelconque de ses points est l'espace quotient de l'espace tangent à la variété ambiante modulo l'espace tangent à la sous-variété au point considéré. Le fibré normal à la sous-variété est la réunion de tous les espaces normaux en tous les points de la sous-variété (munie de la projection naturelle sur cette sous-variété). La solution de l'équation aux variations peut être considérée comme prenant ses valeurs dans le fibré normal au graphe de la solution envisagée de l'équation initiale.

En effet, l'équation aux variations a été définie à l'aide du système de coordonnées (x, y) dans lequel l'axe x était le graphe de la solution considérée. La valeur de la solution de l'équation aux variations en un point est un vecteur de l'y-direction tangent au plan ambiant au point considéré de l'axe x. Sa projection sur l'espace normal à l'axe des x en ce point définit un vecteur du fibré normal. Les solutions de l'équation aux variations définissent ainsi des courbes dans l'espace du fibré normal. *Ces courbes ne dépendent pas du système de coordonnées qui a servi à les construire.* C'est précisément dans ce sens que l'équation aux variations peut être considérée comme une équation dans le fibré normal.

La démonstration de cette assertion résulte immédiatement, par exemple, du lemme prouvé plus haut.

Dans les notations du lemme l'assertion prouvée veut dire que si $F(X, 0) = X$ et Φ_1 est un terme du développement taylorien sur y et p, linéaire en y et p, alors $\hat{\Phi}_1(X, Y, P) = \Phi_1(X, Y, P)$. La dernière égalité est une conséquence immédiate des formules du lemme.

Donc, la structure de plan localement projectif définie par l'équation aux variations est définie sur le fibré normal. ▶

3°. *Transformation de la variable dépendante.*

Soit un difféomorphisme local du plan (x, y) défini par le changement $y = G(X, Z)$, $x = X$, associant au point de coordonnées (x, y) le point de coordonnées (X, Z).

L e m m e. *L'équation*

$$d^2y/dx^2 = \Phi(x, y, p), \quad p = dy/dx$$

se transforme par ce difféomorphisme en l'équation

$$d^2Z/dX^2 = \tilde{\Phi}(X, Z, \Pi), \quad \Pi = dZ/dX,$$

où

$$\tilde{\Phi}(X, Z, \Pi) = \frac{1}{G_Z}[\Phi(X, G)(X, Z), G' + \Pi G_Z) - G'' - 2\Pi G'_Z - \Pi^2 G_{ZZ}];$$

le symbole prime désigne la dérivation partielle par rapport à X, G et ses dérivées sont partout fonctions de X et Z.

◀ Soient $y = v(x)$ solution de l'équation initiale, $Z = V(X)$ son image. On a alors

$$\frac{dv}{dx} = G' + V' G_Z$$

(ici et dans la suite G et ses dérivées sont fonctions de X et $V(X)$). Ensuite

$$\frac{d^2v}{dx^2}=G''+2G'_Z V'+G_{ZZ}V'^2+G_Z V''=\Phi(X, G, G'+XG_Z).$$

En déduisant V'' de la dernière égalité, on obtient la formule du lemme. ▶

Le lemme démontré a pour conséquence immédiate les

Corollaire 1. *Soit* $\Phi=0$. *Alors* $\tilde{\Phi}$ *est un polynôme en* Π *de degré* $\leqslant 2$.

Corollaire 2. *Supposons que* Φ *est un polynôme en* p *de degré* n, $n\geqslant 2$. *Alors* $\tilde{\Phi}$ *est un polynôme en* Π *de degré* $\leqslant n$.

4° *Calcul des termes quadratiques.*

Considérons le difféomorphisme local du plan défini par le changement

$$x = F(X, Z) = X + f(X)Z + O(|Z|^2),$$
$$y = G(X, Z) = Z + g(X)Z^2 + O(|Z|^3).$$

Lemme. *L'image de l'équation*

$$d^2y/dx^2 = \Phi(x, y, p), \ p = dy/dx, \ \Phi = O(|y|^2 + |p|^2)$$

par ce difféomorphisme est

$$d^2Z/dX^2 = \Psi(X, Z, \Pi), \ \Pi = dZ/dX,$$

où

$$\Psi(X, Z, \Pi) = \Phi(X, Z, \Pi) + \Omega(X, Z, \Pi) + O(|Z|^3 + |\Pi|^3),$$
$$\Omega = \alpha Z^2 + \beta Z\Pi + \gamma\Pi^2,$$
$$\alpha = -g'', \ \beta = -4g' + f'', \ \gamma = -2g + 2f'$$

(f, g et leurs dérivées sont fonctions de X).

◀ Ce lemme se démontre par application des lemmes relatifs à la transformation de la variable indépendante d'abord et de la variable dépendante ensuite. En développant les seconds membres des formules obtenues sur (Z, Π) et en ne gardant que les termes quadratiques, on obtient la quantité Ω pour contribution aux termes quadratiques de Φ. ▶

5° *Réduction des termes quadratiques.*

Désignons les termes quadratiques du second membre de Φ par

$$\Phi_2 = Ay^2 + Byp + Cp^2.$$

Les termes quadratiques de l'équation image s'écrivent alors

$$\Psi_2 = (A - g'')Z^2 + (B - 4g' + f'')Z\Pi + (C - 2g + 2f')\Pi^2$$

(les fonctions A, B, C, f, g et leurs dérivées dépendent de X, lequel X coïncide du reste avec x le long de la solution considérée). On vérifie immédiatement le

Lemme. *La quantité*

$$I = 6A - 2B' + C''$$

est invariante par l'application $\Phi_2 \longmapsto \Psi_2$.

En choisissant convenablement les fonctions f et g on peut annuler deux des trois coefficients A, B, C (tout en conservant I). En particulier, si l'on déduit f et g des conditions $4g' - f'' = B$, $2g - 2f' = C$, on obtient $\Psi_2 = \bar{A}Z^2$, $\bar{A} = I/6$, ce qui prouve le théorème.

C. Violation infinitésimale du théorème de Désargues.

Le système de coordonnées dans lequel l'équation différentielle du second ordre s'écrit au voisinage du graphe de l'une de ses solutions sous la forme

$$d^2y/dx^2 = A(x)y^2 + O(|y|^3 + |p|^3), \ p = dy/dx,$$

n'est pas défini de façon unique. Voyons dans quelle mesure le coefficient A dirimant à la rectification de la famille de solutions et qui définit le degré de violation infinitésimale du théorème de Désargues au point donné dans la direction donnée est invariant, c'est-à-dire ne dépend pas du procédé de réduction à la forme normale.

Théorème. *Sur le graphe de la solution nulle la forme différentielle de degré* 5/2

$$\omega = A(x)|dx|^{5/2}$$

est définie intrinsèquement à une constante multiplicative près.

En d'autres termes, si (X, Y) est un autre système de coordonnées dans lequel l'équation s'écrit dans sa forme normale, et si, de plus, la solution $Y = 0$ correspond à $y = 0$ et le coefficient $\overline{A}(X)$ à $A(x)$, alors

$$\overline{A}(X) = CA(x)|dx/dX|^{5/2},$$

où C ne dépend pas de x.

La forme ω définit le *degré de violation du théorème de Desargues sur la solution donnée.*

◀ Le difféomorphisme le plus général laissant invariant l'axe $y = 0$ associe au point (x, y) le point

$$X = f_0(x) + yf_1(x) + \ldots, \quad Y = yg_1(x) + y^2g_2(x) + \ldots$$

L'image du vecteur du fibré normal ayant pour origine le point x et pour y-composante ξ est le vecteur appliqué au point $f_0(x)$ et ayant $g_1(x)\xi$ pour y-composante.

La structure projective du fibré normal est définie intrinsèquement (cf. point B), donc l'application $(x, \xi) \mapsto (f_0(x), g_1(x)\xi)$ construite d'après le difféomorphisme $(x, y) \mapsto (X, Y)$, associant la forme normale d'une équation à la forme normale d'une autre, doit être projective. D'où il suit

$$f_0 = \frac{ax+b}{cx+d}, \quad g_1 = \frac{C}{cx+d}.$$

Tout difféomorphisme laissant invariant l'axe $y = 0$ et préservant la structure projective de son fibré normal peut être représenté sous forme de produit d'une application spéciale

$$X = f_0(x), \ Y = yg_1(x)$$

par un difféomorphisme respectant ponctuellement le fibré normal:

$$(X, Y) \mapsto (X + Yf_1(X) + \ldots, Y + Y^2g_2(X) + \ldots).$$

Le dernier difféomorphisme peut être écrit sous la forme d'un produit des transformations des variables dépendante et indépendante étudiées au point B Donc la quantité I (cf. point B, 5°) reste invariante par ce difféomorphisme.

Etudions le comportement de I par une application projective spéciale.

Toute application projective se décompose en un produit de translations, d'homothéties et de l'application $x \mapsto 1/x$.

La quantité I est invariante par les translations et multipliée par une constante par une homothétie sur x ou y. Il suffit d'étudier son comportement par l'application $x = 1/X$, $Y = Y/X$.

Le calcul des dérivées $P = dY/dX$ et $dP/dX = d^2Y/dX^2$ donne $P = y - px$, $dP/dX = X^{-3}dp/dx$ (où $p = dy/dx$). Donc,

$$d^2Y/dX^2 = X^{-3}A(1/X)(Y/X)^2 + O(|Y|^3 + |P|^3).$$

Par suite, le coefficient en Y^2 est égal à $X^{-5}A(x)$. ▶

D. Construction des invariants scalaires.

La forme différentielle définie plus haut nous permet d'obtenir des fonctions scalaires liées intrinsèquement à l'équation.

On remarquera tout d'abord qu'à une forme différentielle ω (de degré quelconque) définie sur une variété à une dimension est lié intrinsèquement un champ de vecteurs (une forme de degré -1): cette forme prend la valeur 1 sur le vecteur du champ en chaque point.

Par exemple, à la forme $A(x)(dx)^{5/2}$ est lié intrinsèquement le champ $v(x)\partial/\partial x$, où $v = A^{-2/5}$.

Théorème. *Soit $v(x)\,\partial/\partial x$ un champ de vecteurs sur une droite. Alors les fonctions scalaires*

$$I_2 = 2v''v - v'^2,\ I_3 = 2v'''v^2,\ \dots,\ I_n = vI'_{n-1},\ \dots$$

où le symbole prime désigne la dérivation par rapport à x, sont liées intrisèquement au champ relativement aux applications projectives de la droite.

◀ L'invariance de I_2 se vérifie par un calcul immédiat: en effet, il suffit de considérer le changement $x = 1/X$, puisque l'invariance par les translations et les homothéties est évidente. La dérivée de la fonction le long du champ donné est liée à la fonction et au champ intrinsèquement non seulement par rapport aux applications projectives, mais aussi par rapport à tous les difféomorphismes de la droite. Donc l'invariance de tous les I_n résulte de celle de I_2. ▶

Remarque 1. Les considérations suivantes ont présidé à la construction de l'invariant I_2. L'algèbre de Lie du groupe projectif de la droite étant engendrée par les champs $\partial/\partial x$, $x\partial/\partial x$, $x^2\partial/\partial x$ *), tout champ de vecteurs peut, en tout point et aux termes quadratiques inclus près, être approché par un champ projectif (un champ de l'algèbre de Lie du groupe projectif).

L'image du champ projectif approximant le champ de départ en un point par les applications projectives est un champ projectif approximant le champ image au point image. L'action des applications projectives de la droite dans l'espace à trois dimensions des champs projectifs est l'action adjointe du groupe projectif sur son algèbre de Lie. Or, cette action conserve une forme quadratique sur l'algèbre. En effet, si l'on représente les applications projectives par des matrices du second ordre et les champs projectifs par les matrices des opérateurs générateurs des groupes à un paramètre correspondants, l'action de l'application g sur le champ v est représentée par le produit de matrices gvg^{-1}. Or $\det gvg^{-1} = \det v$. Donc le déterminant de la matrice v est une forme quadratique sur l'algèbre de Lie des champs projectifs, invariante par rapport à la représentation adjointe. Par suite, ce déterminant calculé pour le champ projectif approximant le champ de vecteurs donné est une quantité scalaire intrinsèquement liée au champ donné par rapport aux applications projectives.

Le champ projectif approximant possède les composantes $(v, v', v''/2)$ dans la base indiquée plus haut de l'algèbre de Lie. Donc la matrice du champ

---

*) Les groupes à un paramètre d'applications projectives correspondants s'écrivent en coordonnées affines sous la forme $g^t x = x + t$; $g^t x = e^t x$, $g^t x = (x/1 - tx)$ et, par suite, sont définis en coordonnées homogènes par des matrices unimodulaires du second ordre

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},\ \begin{pmatrix} \exp(t/2) & 0 \\ 0 & \exp(-t/2) \end{pmatrix},\ \begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix}.$$

Donc la matrice des opérateurs générateurs est

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}\ \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}.$$

s'écrit

$$\begin{pmatrix} v'/2 & v \\ -v''/2 & -v'/2 \end{pmatrix};$$

et son déterminant n'est autre que I_2 (à un facteur insignifiant près).

Remarque 2. Toute fonction (polynôme, série, . . .) des valeurs de v et d'un nombre fini de dérivées de v liée intrinsèquement au champ v par rapport aux applications projectives s'exprime vraisemblablement au moyen des invariants I_k.

Remarque 3. On peut construire les invariants projectifs d'une fonction sur la droite projective de la manière suivante: on associe à la fonction sa différentielle (une 1-forme), à la forme, un champ, au champ, les invariants I_k. En particulier,

$$I_2[f] = \frac{2f'f''' - 3f''^2}{f'^4}$$

est un invariant élémentaire de la fonction f par rapport aux applications projectives de la variable indépendante. (Cette quantité diffère de la dérivée de Schwartz, qui est invariante par les applications projectives de l'axe des valeurs de la fonction, d'un terme multiplicatif f'^2 au dénominateur.)

Remarque 4. Les invariants I_2, I_3, . . . sont multipliés par λ^2, λ^3, ... si le champ de vecteurs est multiplié par λ. On peut en faire facilement des combinaisons insensibles à la multiplication du champ par un nombre, par exemple, $J = I_2^3/I_3^2$.

La quantité J, correspondant au champ v construit à l'aide de la (5/2)-forme ω, est une fonction scalaire sur l'espace des éléments linéaires du plan, ne dépendant plus du système de coordonnées, mais uniquement de l'équation différentielle initiale du second ordre.

E. Equations cubiques en la dérivée.

La nullité de la forme ω sur toute solution est nécessaire pour la rectification de l'équation (par réduction à la forme $d^2y/dx^2 = 0$) mais, comme nous allons le voir maintenant, pas suffisante.

Théorème 1. *Supposons que l'équation différentielle*

$$d^2y/dx^2 = \Phi\,(x,\ y,\ p),\ p = dy/dx,$$

se ramène à la forme $d^2y/dx^2 = 0$ par un difféomorphisme du plan. Alors Φ est un polynôme en p de degré $\leqslant 3$.

En d'autres termes, dans un système quelconque de coordonnées curvilignes, l'équation différentielle de la famille de toutes les droites du plan est un polynôme en la dérivée première de degré $\leqslant 3$.

Le théorème 1 est une conséquence du fait (étonnant) suivant:

Théorème 2. *Supposons que le second membre de l'équation différentielle*

$$d^2y/dx^2 = \Phi\,(x,\ y,\ p)$$

est un polynôme en p de degré $\leqslant 3$. Alors tout difféomorphisme du plan associe à cette équation une équation de la même forme, c'est-à-dire le second membre reste un polynôme en la dérivée de degré $\leqslant 3$.

◀ Le théorème 2 découle des lemmes du point B relatifs à l'effet des changements des variables indépendante et dépendante sur le second membre de l'équation (cf. corollaire du point B), puisque tout difféomorphisme local du plan est une composition de tels changements. ▶

◀ Le théorème 1 découle du théorème 2 et du fait que 0 est un polynôme en p de degré $\leqslant 3$. ▶

Exercice. Effectuer le changement $(x, y) \to (y, x)$ dans l'équation

$$d^2y/dx^2 = a_0y'^3 - a_1y'^2 + b_1y' - b_0$$

où a_i et b_i sont des fonctions de x et y.

R é p o n s e.

$$d^2x/dy^2 = b_0x'^3 - b_1x'^2 + a_1x' - a_0.$$

R e m a r q u e. On démontre que les conditions $\omega \equiv 0$ et $d^4\Phi/dp^4 \equiv 0$ sont indépendantes. Donc la condition $\omega \equiv 0$ n'est pas suffisante pour ramener l'équation à la forme $d^2y/dx^2 = 0$.

Ces deux conditions $\{\omega \equiv 0,\ d^4\Phi/dp^4 \equiv 0\}$ prises ensemble sont suffisantes pour ramener l'équation à la forme $d^2y/dx^2 = 0$. Ceci ressort d'ailleurs des formules du point B (après quelques calculs).

E x e r c i c e. Montrer que toute équation du second ordre se ramène localement (au voisinage du point $x = 0$ de la solution $y = 0$) à la forme $d^2y/dx^2 = p^2B(x, y, p)$, $p = dy/dx$.

F. Géométrie d'un couple de champs de directions dans l'espace $\mathbb{R}^3$.

Soit un couple de champs de directions dans un espace $\mathbb{R}^3$. On constate que la classification locale de tels couples génériques (aux difféomorphismes de l'espace près) est équivalente à celle des équations différentielles du second ordre (aux difféomorphismes près des variables dépendante et indépendante; localement = au voisinage de l'élément donné d'une droite).

Associons tout d'abord au couple de champs de directions une famille de courbes à deux paramètres sur le plan.

A cet effet, rectifions un champ par un difféomorphisme local de l'espace en transformant la famille des courbes intégrales de ce champ dans la famille de droites parallèles verticales. Ensuite projetons les courbes intégrales du second champ sur un plan horizontal le long des droites verticales. Nous obtenons une famille de courbes à deux paramètres sur un plan horizontal (le plan quotient de l'espace modulo les courbes intégrales du premier champ).

Formons maintenant l'équation différentielle de second ordre dont les graphes des solutions sont les courbes de cette famille à deux paramètres.

On remarquera à cet effet que dans une famille générique à deux paramètres de courbes locales, au voisinage de chaque élément linéaire (points, directions) d'une courbe il passe par chaque point du plan dans chaque direction une courbe de la famille et une seule (il suffit seulement que le jacobien soit non nul).

Si la famille est générique au sens mentionné et (x, y) sont des coordonnées sur le plan, alors la valeur de d^2y/dx^2 sur une courbe de la famille est une fonction différentiable de l'élément considéré, c'est-à-dire de $(x, y, dy/dx)$. Nous obtenons ainsi une équation différentielle du second ordre $d^2y/dx^2 = \Phi(x, y, dy/dx)$. Les graphes des solutions de cette équation font partie de la famille (d'après le théorème d'unicité).

Ainsi, nous avons associé à un couple de champs de directions dans $\mathbb{R}^3$ une équation différentielle du second ordre moyennant une condition de non-dégénérescence: la non-nullité du jacobien.

D é f i n i t i o n. Un champ de directions dans $\mathbb{R}^3$ est dit *régulier* (*non dégénéré*) *par rapport à la verticale* si sa direction n'est nulle part verticale et si la projection horizontale de cette dernière pivote à une vitesse non nulle, lorsque le point d'application se déplace sur une verticale.

Un couple de champs de directions dans $\mathbb{R}^3$ est *régulier* si l'un étant rendu vertical par un difféomorphisme, l'autre devient régulier par rapport au premier.

R e m a r q u e. On voit immédiatement que la définition est correcte. Si le second champ devient régulier par rapport au premier par un difféomorphisme rectifiant les courbes intégrales du premier champ, il le deviendra par tout autre difféomorphisme rectifiant le premier champ. On vérifie immédiatement aussi que la régularité du couple est préservée par une permutation des champs (c'est-à-dire on peut rectifier indifféremment le premier ou le second champ).

Plus haut à tout couple régulier de champs de directions locaux dans $\mathbb{R}^3$ on a associé une équation différentielle du second ordre. En effet, notre condition de régularité coïncide avec la non-nullité du jacobien $\partial(y, y')/\partial(u, v)$, où (u, v) sont les paramètres de la famille. Montrons maintenant que cette cor-

respondance établit une équivalence totale entre les problèmes de la classification locale d'un couple de champs de directions dans $\mathbb{R}^3$ et les équations différentielles du second ordre.

Théorème. *Chaque équation du second ordre s'obtient par la construction précédente à partir d'un couple régulier convenable de champs de directions dans $\mathbb{R}^3$.*

Si deux couples de champs sont associés l'un à l'autre par un difféomorphisme de $\mathbb{R}^3$, alors les équations correspondantes sont associées l'une à l'autre par un difféomorphisme du plan. Inversement, si deux équations du second ordre se correspondent par un difféomorphisme du plan, alors deux couples quelconques de champs leur correspondant sont associés l'un à l'autre par un difféomorphisme de l'espace.

Donc, tout couple de champs est défini de façon univoque (aux difféomorphismes près) de l'espace par l'équation correspondante.

◀ Dans l'espace des 1-jets, muni des coordonnées (x, y, p), considérons la famille des lignes « verticales» de p-direction ($x =$ const, $y =$ const) et la famille des courbes intégrales du système $dy/dx = p$; $dp/dx = \Phi$ équivalent à l'équation $d^2y/dx^2 = \Phi\ (x, y, dy/dx)$.

Une projection verticale nous donne la famille des graphes des solutions de notre équation. Donc, pour chaque équation nous avons exhibé le couple de champs de directions qui l'a déterminée.

Le difféomorphisme de l'espace associant à une droite verticale une droite verticale définit un difféomorphisme du plan horizontal. Donc, si deux couples de champs de $\mathbb{R}^3$ sont difféomorphes, leurs équations correspondantes le seront également.

Inversement, un difféomorphisme du plan horizontal agit sur les éléments linéaires (les directions) des courbes du plan. Donc un difféomorphisme envoyant la première équation dans la seconde définit un difféomorphisme de l'espace $\mathbb{R}^3$ contenant le premier couple de champs de directions dans un autre espace $\mathbb{R}^3$. [L'image d'un point du premier espace est un élément de courbe dans le plan, le difféomorphisme envoie cet élément dans un autre qui est la projection de la direction du second champ du second couple dans un point unique du second espace; et c'est ce point qui est associé au point de départ.] ▶

Ce théorème nous apprend que les résultats obtenus sur la géométrie de la famille des solutions de l'équation du second ordre peuvent être formulés en termes de géométrie d'un couple régulier de champs de directions dans l'espace $\mathbb{R}^3$ (ou encore en termes de géométrie d'un système de deux équations du premier ordre non résolues par rapport aux dérivées dans le cas le plus simple où les dérivées dépendent biunivoquement des coordonnées).

On remarquera également qu'un couple régulier de champs de directions définit dans $\mathbb{R}^3$ une structure de contact (un champ de plans complètement non intégrales, engendrés par ces directions, pour plus de détail, cf. chapitre 2). Les courbes intégrales de nos champs forment des fibrés *) legendriens au sens de cette structure. Nous pouvons donc reformuler les résultats précédents en termes de géométrie de couples de fibrés legendriens dans $\mathbb{R}^3$.

E. Dualité.

Plus haut nous avons construit pour chaque couple régulier de champs de directions dans $\mathbb{R}^3$ une équation différentielle du second ordre en rectifiant un champ et en projetant les courbes intégrales de l'autre. On pouvait cependant procéder de façon inverse, c'est-à-dire rectifier le second champ et projeter les courbes intégrales du premier. En définitive, on aurait obtenu une autre équation du second ordre.

Donc, toute équation différentielle du second ordre possède une équation duale.

---

*) On appelle variété legendrienne d'une structure de contact la sous-variété intégrale de dimension maximale. Un fibré legendrien est un fibré de fibres legendriennes.

En termes de couples de champs de directions dans $\mathbb{R}^3$, le passage à l'équation duale se traduit par une permutation des champs du couple. Voici une autre description de l'équation duale. Supposons que la famille de solutions des équations différentielles du second ordre en $y(x)$ dépendant des paramètres (u, v) est écrite sous la forme $F(x, y; u, v) = 0$. Supposons maintenant que x et y sont les paramètres et u et v, les variables. La même relation définit une famille à deux paramètres de fonctions $v(u)$. Cette famille est une famille de solutions d'une équation du second ordre qui n'est autre que l'équation duale à l'initiale.

Exercice. Prouver que la forme ω (cf. point B) est nulle si et seulement si le second membre est un polynôme en la dérivée de degré $\leqslant 3$.

Donc, la condition de rectification des équations du second ordre peut s'énoncer :

L'équation $d^2y/dx^2 = \Phi(x, y, dy/dx)$ se ramène à la forme $d^2y/dx^2 = 0$ si et seulement si cette équation et sa duale ont pour second membre un polynôme en la dérivée de degré $\leqslant 3$.

F. Aperçu.

La géométrie de l'équation du second ordre a inspiré de nombreuses théories mathématiques.

1° A. Tressé, élève de Lie, a construit tous les semi-invariants de l'équation dans sa thèse « Détermination des invariants ponctuels de l'équation différentielle ordinaire de second ordre », Leipzig, 1896 (voir également son article « Sur les invariants différentiels des groupes continus des transformations », Acta mathematica, 1894, 18, 1-88).

[On appelle semi-invariant d'ordre k une fonction de la valeur prise par le second membre de l'équation et par ses dérivées d'ordre $\leqslant k$ en un point donné de l'espace (x, y, p), dont la nullité en ce point entraîne celle de la valeur prise par cette fonction sur l'image de l'équation par un difféomorphisme du plan au point image.]

Il existe très exactement deux semi-invariants fonctionnellement indépendants d'ordre quatre. L'un d'eux est $\partial^4\Phi/\partial p^4$. L'autre est l'invariant scalaire I_2 construit d'après la forme ω (cf. point D). Pour l'équation

$$d^2y/dx^2 = A(x)\, y^2 + O(|y|^3 + |p|^3),$$

on a

$$I_2 = 2v''v - v'^2, \quad v = A^{-2/5};$$

donc le semi-invariant est $5AA'' - 12A'^2$.

Tressé a trouvé encore trois semi-invariants d'ordre cinq et pour $k > 5$ encore $(k^2 - k - 8)/2$ semi-invariants d'ordre k; les autres semi-invariants sont fonctions des précédents. Tressé a encore signalé que tous les semi-invariants se déduisent par dérivation des trois semi-invariants élémentaires.

2°. Le problème de la géométrie de l'équation différentielle de second ordre a conduit Elie Cartan à la théorie des variétés à connexion projective (E. Cartan, Sur les variétés à connexion projective, Bull. Soc. Math. France, 52, 1924, 205-241; Œuvres, III, 1, n° 70, Paris, 1955, 825-862).

La connexion projective sur une variété consiste à associer à un chemin différentiable sur cette variété une application projective dépendant différentiablement du chemin *) de l'espace tangent à la variété en l'origine du chemin dans l'espace tangent en l'extrémité. En particulier, à un chemin fermé est associée une application projective de l'espace tangent initial, et au parcours d'un « parallélogramme infinitésimal », une « application projective infinitésimale » appelée forme de courbure de la connexion.

*) Vérifiant des conditions naturelles que nous n'exhiberons pas.

La connexion projective est dite connexion sans torsion si l'origine des coordonnées de l'espace tangent reste fixe par un transport sur un chemin fermé. Parmi les connexions sans torsion Cartan distingue les connexions normales. En dimension deux la normalité de la connexion signifie que toute droite issue de l'origine des coordonnées du plan tangent est associé à elle-même par un déplacement le long d'un chemin fermé.

On appelle ligne géodésique d'une connexion une ligne d'une variété dont la tangente se transforme en une tangente par déplacement le long de cette ligne.

Les géodésiques d'une connexion projective sans torsion sur le plan sont graphes des solutions d'une équation du second ordre dont le second membre est un polynôme en la dérivée de degré $\leqslant 3$. Inversement, à toute équation du second ordre dont le second membre est un polynôme en la dérivée de degré $\leqslant 3$ est associée exactement une connexion projective normale sans torsion pour laquelle les géodésiques sont graphes.

S'agissant des équations générales du second ordre Cartan leur associe de façon univoque une connexion projective normale sans torsion, le déplacement du plan ayant lieu sur des chemins de l'espace à trois dimensions des éléments (pour les détails voir le travail cité plus haut).

3°. G. Bol (Über topologische Invarianten von zwei Kurvenscharen in Raum, Abhandlungen Math. Sem. Univ. Hamburg, 9, 1 (1932), 15-47) a traduit la théorie de Tressé dans le langage d'un couple de champs de directions dans l'espace.

Cette théorie laisse semble-t-il entier le problème de l'espace des orbites, et, en particulier, le problème du nombre de paramètres numérotant les orbites des k-jets de l'équation du second ordre (sous l'action du groupe de difféomorphismes du plan dans l'espace des k-jets des équations) au voisinage d'un point générique de l'espace des k-jets.

CHAPITRE 2

EQUATIONS AUX DÉRIVÉES PARTIELLES DU PREMIER ORDRE

Les équations aux dérivées partielles sont bien moins étudiées que les équations différentielles ordinaires. Seule la théorie de l'équation à dérivées partielles du premier ordre se ramène à l'étude d'équations différentielles ordinaires spéciales appelées équations des caractéristiques. Ce lien est dû au fait qu'un milieu continu constitué de particules indépendantes est susceptible d'être décrit aussi bien par une équation différentielle aux dérivées partielles pour un champ que par des équations différentielles ordinaires pour des particules.

Dans ce chapitre on se propose d'exposer cette théorie (qui mathématiquement se ramène à la géométrie des structures de contact). On étudie également l'intégrabilité d'un champ d'hyperplans (théorème de Frobénius).

§ 7. Equations linéaires et quasi linéaires à dérivées partielles du premier ordre

L'intégration de l'équation à dérivées partielles du premier ordre se ramène à celle d'un système d'équations différentielles ordinaires dites équations des caractéristiques. Cette réduction est basée sur des considérations géométriques simples relatives à la formation de surfaces à partir de familles de courbes. Nous développerons d'abord ces généralités géométriques et ensuite nous les appliquerons aux équations aux dérivées partielles.

A. Surfaces intégrales d'un champ de directions.

Soient X une variété différentiable *), V un champ de directions sur X.

D é f i n i t i o n. Une sous-variété différentiable $Y \subset X$ est une surface intégrale du champ V si tout plan tangent à Y contient une direction du champ V (fig. 39).

*) On dira variété différentiable pour indéfiniment différentiable.

Théorème. *Une sous-variété $Y \subset X$ est une surface intégrale du champ V si et seulement si avec chacun de ses points elle contient un intervalle de courbe intégrale passant par ce point.*

◀ Cette assertion est locale et invariante par rapport aux difféomorphismes. Il suffit donc de la prouver pour un champ standard de directions parallèles dans un espace vectoriel. Or dans ce cas

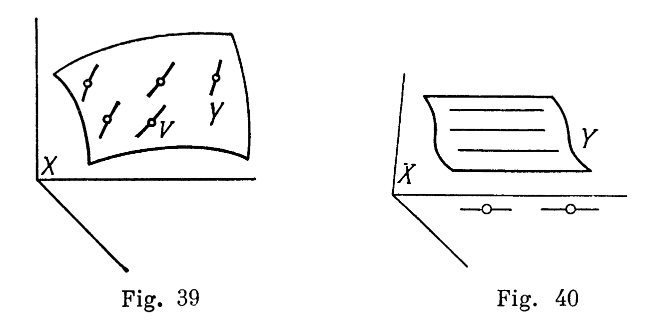


Fig. 39 Fig. 40

cette proposition est évidente (elle revient à affirmer qu'une fonction définie sur un intervalle est constante si et seulement si sa dérivée est nulle, fig. 40). ▶

Soit Γ une sous-variété de dimension k d'une variété X de dimension n (fig. 41). Γ est une hypersurface si $k = n - 1$.

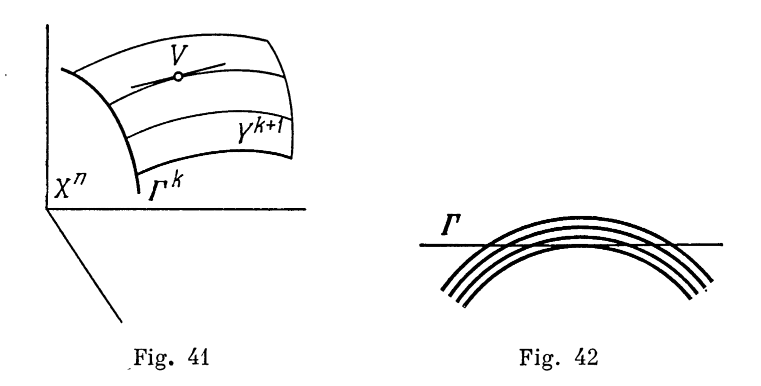


Fig. 41 Fig. 42

Définition. On appelle *problème de Cauchy* pour un champ de directions V de variété initiale Γ le problème qui consiste à chercher une sous-variété intégrale de dimension $k + 1$ du champ V renfermant la sous-variété initiale Γ.

On remarquera que les courbes intégrales passant par Γ ne forment pas toujours une sous-variété même localement au voisinage de la variété Γ (cf. fig. 42).

D é f i n i t i o n. On appelle point *caractéristique* dans le champ de directions V le point de tangence d'une direction de ce champ à la variété initiale.

T h é o r è m e. *Soit donné un point non caractéristique d'une variété initiale de dimension k dans un champ de directions V. Il existe alors une surface intégrale de dimension $k+1$ de ce champ contenant un voisinage de ce point. Cette surface est unique en ce sens que deux surfaces intégrales quelconques contenant les voisinages du point indiqué sur la variété initiale sont confondues dans un voisinage de ce point.*

◀ Cette proposition est vraie localement et invariante par les difféomorphismes. Il suffit donc de la prouver pour un champ standard de directions parallèles dans un espace vectoriel. Dans ce cas la proposition est évidente (pourquoi?). ▶

B. Equation homogène linéaire de premier ordre.

D é f i n i t i o n. On appelle *équation homogène linéaire de premier ordre* l'équation

$$L_a u = 0, \tag{1}$$

où a est un champ de vecteurs défini sur une variété M.

Soient $(a_1, \ldots, a_n)$ les composantes du champ a dans les coordonnées $(x_1, \ldots, x_n)$. Chaque composante est une fonction de $(x_1, \ldots, x_n)$. L'équation (1) s'écrit

$$a_1 \frac{\partial u}{\partial x_1} + \ldots + a_n \frac{\partial u}{\partial x_n} = 0. \tag{2}$$

D é f i n i t i o n. Le champ a s'appelle *champ de vecteurs caractéristique* de l'équation (1), ses trajectoires, *caractéristiques*, l'équation $\dot{x} = a(x)$, *équation des caractéristiques* pour l'équation aux dérivées partielles (1).

T h é o r è m e. *Une fonction u est solution de l'équation linéaire du premier ordre si et seulement si elle est intégrale première de l'équation des caractéristiques.*

◀ On reconnaît ici la définition des intégrales premières. ▶

D é f i n i t i o n. On appelle *problème de Cauchy* pour l'équation (1) le problème qui consiste à trouver la solution u vérifiant la condition $u|_\gamma = \varphi$, où γ est une hypersurface différentiable dans M, et φ une fonction différentiable donnée sur cette hypersurface.

L'hypersurface γ est appelée *hypersurface initiale*, la fonction φ, condition initiale.

A noter que le problème de Cauchy n'admet pas toujours de solution. En effet, la solution u est constante le long de chaque caractéristique. Or la caractéristique est susceptible de couper plusieurs fois la surface γ. Si la fonction donnée φ prend des valeurs distinctes en ces points, le problème de Cauchy ne possède aucune solution

dans aucun domaine renfermant la caractéristique mentionnée (fig. 43).

D é f i n i t i o n. On dit qu'un point x de l'hypersurface initiale γ est *non caractéristique* si la caractéristique passant par ce point est transversale (non tangente) à γ.

T h é o r è m e. *Soit x un point non caractéristique de γ. Il existe alors un voisinage du point x tel que le problème de Cauchy pour l'équation* (1) *possède une solution et une seule dans ce voisinage.*

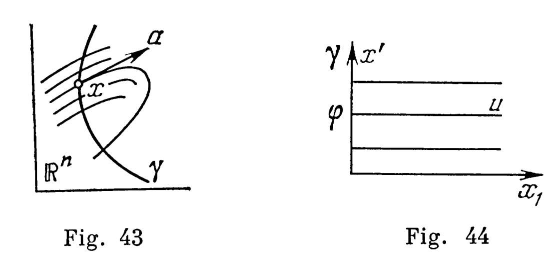


Fig. 43 Fig. 44

◀ Le théorème de rectification nous dit qu'on peut choisir au voisinage du point x des coordonnées telles que les composantes du champ a soient (1, 0, ..., 0) et l'équation de γ prenne la forme $x_1 = 0$ (fig. 44).

Dans ces coordonnées le problème de Cauchy devient

$$\frac{\partial u}{\partial x_1} = 0, \quad u|_{x_1=0} = \varphi. \tag{3}$$

L'unique solution (dans un domaine convexe) est: $u(x_1, x') = \varphi(x')$, où $x' = (x_2, \ldots, x_n)$. ▶

R e m a r q u e. Les solutions de toute équation différentielle ordinaire forment une variété finidimensionnelle: chaque solution est définie par une collection finie de conditions initiales. On voit que l'équation aux dérivées partielles homogène linéaire de premier ordre en une fonction de n variables possède « autant de solutions qu'il n'existe de fonctions de $n - 1$ variables ».

On verra par ailleurs que ceci est également vrai pour les équations générales aux dérivées partielles du premier ordre.

C. Equation non homogène linéaire du premier ordre.

D é f i n i t i o n. On appelle équation non homogène linéaire du premier ordre l'équation

$$L_a u = b, \tag{4}$$

où a est un champ de vecteurs défini sur une variété M, b une fonction définie sur M. L'équation (4) s'écrit encore

$$a_1 \frac{\partial u}{\partial x_1} + \ldots + a_n \frac{\partial u}{\partial x_n} = b, \tag{5}$$

où $a_k = a_k(x_1, \ldots, x_n)$, $b = b(x_1, \ldots, x_n)$.

Le problème de Cauchy se pose comme pour l'équation homogène (1).

T h é o r è m e. *Le problème de Cauchy pour l'équation* (4) *possède une solution unique dans un voisinage suffisamment petit de tout point non caractéristique* x_0 *de la surface initiale* γ; *cette solution est définie par*

$$u(g(x, t)) = \varphi(x) + \int_0^t b(g(x, \tau))\, d\tau,$$

où $g(x, t)$ *est la valeur prise par la solution de l'équation des caractéristiques (vérifiant la condition initiale* $g(x, 0) = x$ *sur la surface initiale) à l'instant* t.

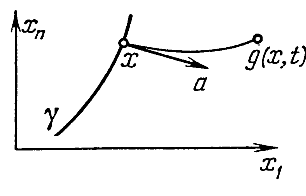


Fig. 45

◀ Après rectification du champ a on est conduit au problème

$$\frac{\partial u}{\partial x_1} = b, \quad u_{x_1=0} = \varphi(x'),$$

dont l'unique solution est

$$u(x_1, x') = \varphi(x') + \int_0^{x_1} b(\xi, x')\, d\xi. \ ▶$$

En d'autres termes, l'équation (4) signifie que la dérivée de la solution le long des caractéristiques est la fonction connue b. Donc, l'accroissement de la solution sur un intervalle de caractéristique est égal à l'intégrale de la fonction b sur cet intervalle (fig. 45).

D. Equation quasi linéaire du premier ordre.

D é f i n i t i o n. On appelle *équation quasi linéaire du premier ordre* l'équation

$$L_\alpha u = \beta, \tag{6}$$

où $\alpha(x) = a(x, u(x))$, $\beta(x) = b(x, u(x))$; x est un point d'une variété M, u une fonction inconnue définie sur M, a un champ de vecteurs donné, tangent à M, dépendant du point $u \in \mathbb{R}$ comme d'un paramètre, b une fonction définie sur $M \times \mathbb{R} = J^0(M, \mathbb{R})$.

Dans les coordonnées $(x_1, \ldots, x_n)$ l'équation (6) s'écrit

$$a_1(x, u)\frac{\partial u}{\partial x_1} + \ldots + a_n(x, u)\frac{\partial u}{\partial x_n} = b(x, u). \tag{7}$$

La seule distinction de l'équation linéaire est que les coefficients a_k et b sont susceptibles de dépendre de la fonction inconnue.

E x e m p l e. Considérons un milieu de dimension un composé de particules se déplaçant chacune par inertie sur une droite à une vitesse constante. Désignons par $u(x, t)$ la vitesse d'une particule au point x à l'instant t. Ecrivons l'équation de Newton: l'accélération de la particule est nulle. Si $x = \varphi(t)$ représente le chemin parcouru par la particule, alors

$$\dot{\varphi} = u(\varphi(t), t)$$

et

$$\ddot{\varphi} = \frac{\partial u}{\partial x}\dot{\varphi} + \frac{\partial u}{\partial t} = u\frac{\partial u}{\partial x} + \frac{\partial u}{\partial t}.$$

Donc, le champ de vitesses u d'un milieu composé de particules indépendantes vérifie l'équation quasi linéaire

$$uu_x + u_t = 0. \tag{8}$$

E x e r c i c e. Tracer le graphe de la fonction $u(\cdot, t)$ sachant que celui de la fonction $u(\cdot, 0)$ est représenté sur la figure 46.

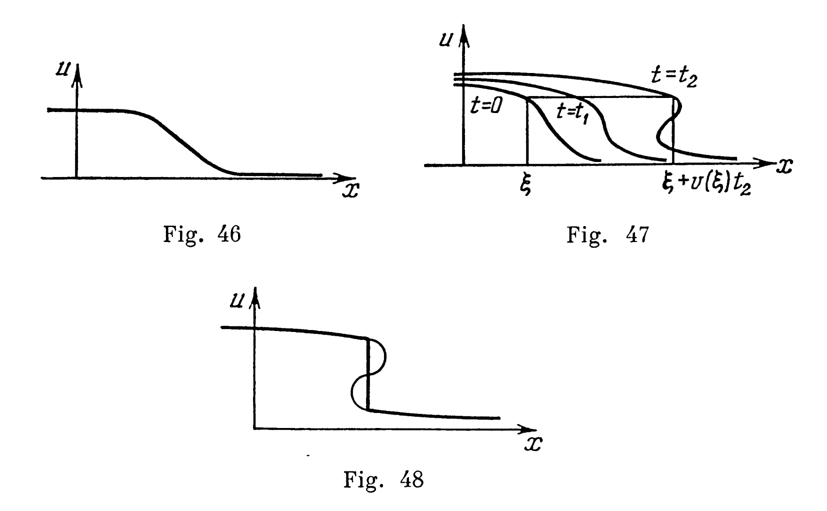


Fig. 46

Fig. 47

Fig. 48

R é p o n s e. Voir fig. 47. Il n'existe pas de solution différentiable pour $t \geqslant t_1$. A partir de cet instant les particules se heurtent dans le milieu. [La condition physique de mouvement par inertie, c'est-à-dire l'absence d'interaction entre les particules, devient irréelle et doit être remplacée par une autre condition: la description du caractère des chocs. On se trouve ainsi en pré-

sence d'ondes de choc qui sont des fonctions de la forme représentée sur la figure 48 vérifiant l'équation (8) en dehors du saut et des conditions supplémentaires de nature physique sur le saut.]

E. Caractéristiques de l'équation quasi linéaire du premier ordre. Nous venons juste de voir l'intérêt qu'il y a de passer du champ des vitesses au mouvement des particules pour l'équation quasi linéaire spéciale (8). Ceci vaut également pour l'équation générale (6). Les mouvements des particules seront alors des courbes dans le produit cartésien du domaine de définition par le domaine de valeurs de la fonction inconnue; ces courbes sont appelées caractéristiques de l'équation quasi linéaire.

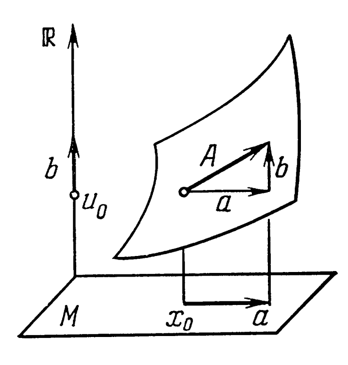


Fig. 49

L'équation quasi linéaire (6) en la fonction inconnue $u: M \to \mathbb{R}$

$$L_{a(x,\,u(x))}u = b\,(x,\,u\,(x)) \qquad (6)$$

traduit le fait que le point x commence à se déplacer à partir du point x_0 dans M à la vitesse $a\,(x_0,\,u_0)$, la valeur de la solution $u = u_0$ se met à varier à la vitesse $b\,(x_0,\,u_0)$.

En d'autres termes, le vecteur $A\,(x_0,\,u_0)$ appliqué au point $(x_0,\,u_0)$ de l'espace $M \times \mathbb{R}$ et de composantes $a\,(x_0,\,u_0)$ sur M et $b\,(x_0,\,u_0)$ sur $\mathbb{R}$ est tangent au graphe de la solution (fig. 49).

D é f i n i t i o n. Le vecteur $A\,(x_0,\,u_0)$ s'appelle *vecteur caractéristique* de l'équation quasi linéaire (6) au point $(x_0,\,u_0)$. Les vecteurs caractéristiques appliqués en tous les points de l'espace $M \times \mathbb{R}$ forment un *champ de vecteurs A* appelé *champ de vecteurs caractéristique* de l'équation quasi linéaire (6). Les trajectoires du champ de vecteurs caractéristique s'appellent *caractéristiques* de l'équation quasi linéaire.

L'équation différentielle définie par le champ de la vitesse de phase de A s'appelle *équation des caractéristiques.*

E x e m p l e. Supposons que M est l'espace $\mathbb{R}^n$ muni des coordonnées $(x_1,\,\ldots,\,x_n)$. Le champ caractéristique est défini par ses composantes; au point $(x,\,u)$ elles sont égales à

$$a_1\,(x,\,u),\;\ldots,\;a_n\,(x,\,u);\;b\,(x,\,u).$$

L'équation des caractéristiques s'écrit

$$\dot{x}_1 = a_1\,(x,\,u),\;\ldots,\;\dot{x}_n = a_n\,(x,\,u);\;\dot{u} = b\,(x,\,u).$$

E x e r c i c e. Trouver les caractéristiques de l'équation d'un milieu composé de particules indépendantes $uu_x + u_t = 0$.

R é p o n s e. $\dot{x} = u$, $\dot{t} = 1$, $\dot{u} = 0$. Les caractéristiques sont les droites $x = a + bt$, $u = b$.

R e m a r q u e. L'équation linéaire est un cas particulier de l'équation quasi linéaire. Cependant les caractéristiques de l'équation linéaire ne sont pas les mêmes que celles de la même équation considérée comme quasi linéaire: les caractéristiques de l'équation linéaire sont contenues dans M, celles de l'équation quasi linéaire dans $M \times \mathbb{R}$. Les caractéristiques de l'équation linéaire sont les projections de celles de la même équation considérée comme quasi linéaire de $M \times \mathbb{R}$ sur M.

F. Intégration de l'équation quasi linéaire de premier ordre. Soit A champ de vecteurs caractéristique de l'équation quasi linéaire (6). Supposons que ce champ est partout non nul. Il définit donc un champ de directions.

D é f i n i t i o n. Le champ de directions du champ de vecteurs caractéristique de l'équation caractéristique s'appelle *champ de directions caractéristique* de cette équation.

Les caractéristiques de l'équation quasi linéaire sont les courbes intégrales du champ de directions caractéristique.

E x e m p l e. Dans le cas, où M est l'espace $\mathbb{R}^n$ muni des coordonnées $(x_1, \ldots, x_n)$, l'équation des caractéristiques s'écrit ordinairement *sous la forme dite symétrique*

$$\frac{dx_1}{a_1} = \frac{dx_2}{a_2} = \ldots = \frac{dx_n}{a_n} = \frac{du}{b}$$

qui traduit la colinéarité de la tangente à la caractéristique au vecteur caractéristique.

T h é o r è m e. *Une condition nécessaire et suffisante pour qu'une fonction soit solution de l'équation quasi linéaire est que son graphe soit surface intégrale du champ de directions caractéristique.*

◀ Ceci est évident, puisque l'équation (6) traduit la tangence du graphe au vecteur caractéristique. ▶

C o r o l l a i r e. *Une condition nécessaire et suffisante pour qu'une fonction soit solution de l'équation quasi linéaire est que son graphe contienne avec chacun de ses points un intervalle de caractéristique passant par ce point.*

◀ Voir point A. ▶

Intégrer l'équation quasi linéaire revient donc à en chercher les caractéristiques. Si les caractéristiques sont connues, il reste seulement à en composer une surface qui est le graphe d'une fonction. Cette fonction sera solution de l'équation quasi linéaire; toutes les solutions s'obtiennent de cette façon.

G. Problème de Cauchy pour l'équation quasi linéaire de premier ordre.

Soient $\gamma \subset M$ une hypersurface (sous-variété de codimension 1) dans M et $\varphi: \gamma \to \mathbb{R}$ une fonction différentiable (fig. 50).

D é f i n i t i o n. On appelle *problème de Cauchy* pour l'équation quasi linéaire (6) *avec la condition initiale* φ *sur* γ, le problème consistant à trouver la solution u qui s'identifie à φ sur γ.

Ce problème se ramène aisément à un problème de Cauchy pour un champ de directions caractéristique.

Considérons le graphe de la fonction $\varphi: \gamma \to \mathbb{R}$. Ce graphe est une hypersurface dans le produit direct $\gamma \times \mathbb{R}$. Comme γ est plongée dans M nous pouvons considérer le graphe Γ de la fonction φ comme une sous-variété de codimension 2 dans $M \times \mathbb{R}$ (fig. 50).

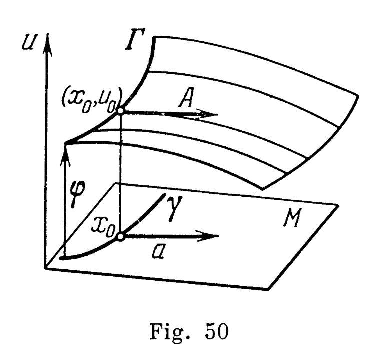


Fig. 50

D é f i n i t i o n. *On appelle sous-variété initiale* pour la condition initiale φ sur γ la sous-variété $\Gamma \subset M \times \mathbb{R}$, graphe de la fonction φ sur γ.

La sous-variété initiale Γ définit donc aussi bien l'hypersurface γ dans M que la condition initiale φ sur γ.

D é f i n i t i o n. La condition initiale (γ, φ) est dite *non caractéristique pour l'équation quasi linéaire* (6) *en un point* x_0 *de* γ si le vecteur $a(x_0, u_0)$ $(u_0 = \varphi(x_0))$ n'est pas tangent à la surface γ en ce point (fig. 50).

[R e m a r q u e. Si l'équation est linéaire, le vecteur $a(x_0, u_0)$ ne dépend pas de u_0 et l'on peut donc définir les points non caractéristiques de la surface γ. Pour l'équation quasi linéaire seuls les points $(x_0, u_0) \in \Gamma$ sont caractéristiques ou non caractéristiques, tandis que pour les points $x_0 \in \gamma$ être caractéristiques n'a aucun sens.]

T h é o r è m e. *L'équation quasi linéaire* (6) *avec une condition initiale non caractéristique au point* x_0 *admet une solution localement unique au voisinage de ce point.*

◀ La condition initiale étant non caractéristique au point x_0, il suit que:

1) Le champ caractéristique A ne s'annule pas au voisinage du point (x_0, u_0). Donc au voisinage de ce point est défini un champ différentiable de directions caractéristiques.

2) La direction caractéristique n'est pas tangente à la variété initiale Γ au point considéré et, partant, à son voisinage. Donc il existe une surface intégrale localement unique du champ caractéristique de directions contenant la variété initiale Γ (cf. point A).

3) Le plan tangent à la surface intégrale au point (x_0, u_0) n'est pas vertical (ne contient pas l'axe u). Donc la surface intégrale est graphe d'une fonction qui n'est autre que la solution cherchée (voir point F). ▶

R e m a r q u e. La démonstration indique également une solution du problème de Cauchy pour l'équation quasi linéaire.

§ 8. Equation non linéaire aux dérivées partielles de premier ordre

Les équations non linéaires aux dérivées partielles de premier ordre s'intègrent à l'aide des caractéristiques comme les équations linéaires. Mais si les caractéristiques de l'équation linéaire en une fonction sur M sont contenues dans M et celles de l'équation quasi linéaire dans $M \times \mathbb{R}$, les caractéristiques de l'équation générale non linéaire aux dérivées partielles de premier ordre appartiennent à la variété des 1-jets des fonctions $J^1(M, \mathbb{R})$.

La variété des 1-jets des fonctions possède une structure de contact naturelle.

L'intégration des équations non linéaires aux dérivées partielles de premier ordre est basée sur des faits simples de la géométrie de la structure de contact par lesquels nous allons entamer notre exposé.

A. Variétés de contact.

On appelle variété de contact une variété munie d'un champ d'hyperplans dans les espaces tangents et vérifiant une condition de « non-intégrabilité maximale ».

Le champ de plans (contrairement au champ de directions à une dimension) peut ne pas posséder de surface intégrale de dimension égale à celle des plans. Pour mesurer l'obstacle à l'existence d'hypersurfaces intégrales du champ d'hyperplans on procèdera à la construction suivante.

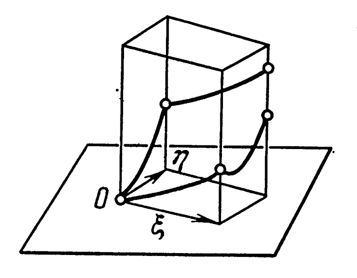


Fig. 51

Soit un point O de la variété. Introduisons au voisinage de ce point des coordonnées telles que le plan du champ passant par O soit hyperplan de coordonnées; on appellera *horizontales* les coordonnées correspondantes, et *vertical* l'axe de coordonnées restant.

Construisons un cylindre sur tout chemin du plan horizontal de génératrice verticale. La trace de notre champ de plans sur la surface du cylindre est un champ de directions. Les surfaces intégrales, si elles existent, coupent les cylindres suivant les courbes intégrales de ce champ de directions (fig. 51).

On peut donc relever tout chemin du plan horizontal sur la surface cherchée.

Soient maintenant (ξ, η) deux vecteurs du plan de coordonnées horizontal passant par le point considéré. Etudions le parallélogramme construit sur (ξ, η). Deux chemins suivant les côtés mènent d'un sommet au sommet opposé. En élevant ces chemins on obtient généralement deux points distincts au-dessus du sommet opposé.

La non-coïncidence de ces points est précisément l'obstacle à la construction d'une hypersurface intégrale, ou encore, à l'« intégrabilité » du champ d'hyperplans.

Considérons la différence des coordonnées verticales des deux points obtenus. La partie principale bilinéaire (en ξ et η) de cette différence mesure le degré de non-intégrabilité du champ. Pour en donner une définition formelle nous allons effectuer la construction suivante.

Un champ d'hyperplans tangents sur la variété est définie localement par une 1-forme différentielle α partout non nulle et définie à la multiplication près par une fonction non nulle : les plans du champ sont les zéros de la forme (les sous-espaces de l'espace tangent sur lesquels la forme est nulle).

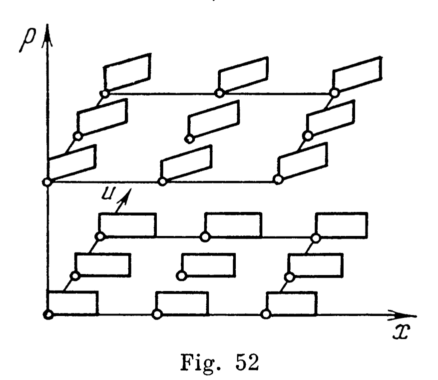


Fig. 52

E x e m p l e. Considérons dans $\mathbb{R}^{2n+1}$ muni des coordonnées $(x_1, \ldots, x_n;$ $u;$ $p_1, \ldots, p_n)$ la 1-forme $\alpha = du - p\,dx$ (où $p\,dx = p_1 dx_1 + \ldots + p_n dx_n$). La 1-forme α ne s'annule en aucun point de $\mathbb{R}^{2n+1}$, donc, elle définit un champ de plans $\alpha = 0$ de dimension $2n$ dans $\mathbb{R}^{2n+1}$ (fig. 52).

D é f i n i t i o n. Une 1-forme différentielle α partout non nulle sur M est dite *de contact* si sa dérivée extérieure $d\alpha$ définit dans chaque plan $\alpha = 0$ une 2-forme extérieure non dégénérée.

[Une forme bilinéaire $\omega: L \times L \to \mathbb{R}$ est non dégénérée si $\forall\, \xi \in L \setminus 0\ \exists\, \eta \in L:\ \omega(\xi, \eta) \neq 0$.]

Une 2-forme extérieure non dégénérée sur un espace vectoriel est dite aussi *structure symplectique.*

E x e m p l e. La forme construite dans l'exemple précédent est de contact. En effet, la dérivée extérieure de la forme α vaut

$$dx_1 \wedge dp_1 + \ldots + dx_n \wedge dp_n.$$

Le plan $\alpha = 0$ est rapporté aux coordonnées $(x_1, \ldots, x_n;\ p_1, \ldots, p_n)$. Dans ces coordonnées la matrice de la forme $\omega = d\alpha\,|_{\alpha=0}$ s'écrit $\begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}$, où E est la matrice unité d'ordre n. Le déterminant de cette matrice étant égal à l'unité, la 2-forme est non dégénérée.

R e m a r q u e. Les formes bilinéaires extérieures non dégénérées n'existent que dans les espaces de *dimension paire*. Donc les formes de contact n'existent que sur les variétés *de dimension impaire*.

T h é o r è m e. *Soient α forme de contact, f fonction non nulle partout. Alors $f\alpha$ est aussi une forme de contact. De plus, les structures symplectiques*

$$d\alpha|_{\alpha=0},\quad d(f\alpha)|_{f\alpha=0}$$

coïncident sur le plan $\alpha = 0$ à un facteur multiplicatif non nul près.

◀ La formule de Leibniz dit que

$$d(f\alpha) = df \wedge \alpha + fd\alpha.$$

Or $df \wedge \alpha$ est une 2-forme nulle sur le plan $\alpha=0$. Donc, les 2-formes $d\alpha$ et $d(f\alpha)$ coïncident sur le plan $\alpha = 0$ à un facteur multiplicatif non nul f. En particulier, la 2-forme $d(f\alpha)|_{\alpha=0} = fd\alpha\,|_{\alpha=0}$ n'est pas dégénérée, par suite, $f\alpha$ est une forme de contact. ▶

D é f i n i t i o n. On appelle *structure de contact* sur une variété M un champ d'hyperplans tangents localement défini comme l'ensemble des zéros d'une 1-forme de contact. Les hyperplans du champ sont appelés *hyperplans de contact*. On désignera par Π_x l'hyperplan de contact au point x (fig. 53).

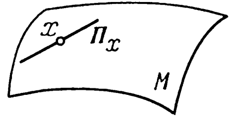


Fig. 53

R e m a r q u e 1. Le théorème précédent nous apprend que la propriété d'une 1-forme définissant un champ de plans d'être de contact ne dépend pas du choix de la forme mais du champ de plans de contact. En effet, si β est une autre forme définissant le même champ, alors β coïncide (localement) avec α à un facteur multiplicatif non nul près. Donc les formes β et α sont simultanément de contact ou de non-contact.

R e m a r q u e 2. Les structures de contact n'existent que sur les variétés de dimension impaire.

D é f i n i t i o n. Une sous-variété d'une variété de contact est une *variété intégrale d'un champ de plans de contact* si tout plan tangent à la sous-variété appartient à un plan de contact.

E x e r c i c e. Prouver que *la dimension d'une variété intégrale d'un champ de plans de contact sur une variété de contact de dimension $2n + 1$ est* $\leqslant n$.

S o l u t i o n. Sur une telle variété Y la forme $i^*\alpha$ (où $i\colon Y \to M^{2n+1}$ est un plongement) est nulle. Donc $i^*d\alpha = di^*\alpha = 0$. Par suite, les vecteurs de l'espace tangent sont deux à deux gauches orthogonaux: $\omega(\xi, \eta) = 0$. D'où il suit que la dimension de l'espace est $\leqslant n$ (voir plus bas point D, problème 2).

R e m a r q u e. Le champ de contact défini sur M^{2n+1} possède des variétés intégrales de dimension n. On les appelle *sous-variétés*

legendriennes. Nous allons voir qu'à toute fonction est associée une sous-variété legendrienne dans l'espace des 1-jets.

B. Structure de contact sur une variété de 1-jets de fonctions. Soient V variété de dimension n, $J^1(V, \mathbb{R})$ la variété des 1-jets des fonctions sur V.

Un 1-jet d'une fonction f sur V est défini par un point $x \in V$, la valeur $u = f(x)$ de la fonction f en x et la première différentielle de la fonction f en x. Donc, la variété des 1-jets des fonctions sur V est de dimension $2n + 1$. Si $(x_1, \ldots, x_n)$ sont des coordonnées locales sur V, alors un 1-jet d'une fonction f sur V est défini par une collection de $2n + 1$ nombres $(x_1, \ldots, x_n ; u ; p_1, \ldots, p_n)$, où $p_i = (\partial f/\partial x_i)(x)$.

Définition. On appelle *forme de contact standard* sur $J^1(V, \mathbb{R})$ une variété des 1-jets des fonctions sur V la 1-forme

$$\alpha = du - pdx \quad (pdx = p_1dx_1 + \ldots + p_ndx_n).$$

Nous avons vu plus haut que c'est une 1-forme de contact sur $\mathbb{R}^{2n+1}$.

Il est aisé de voir que la forme α ne dépend pas du choix des coordonnées $(x_1, \ldots, x_n)$ qui ont servi à la définir et qu'elle est définie globalement.

Définition. On appelle 1-*graphe* d'une fonction $f: V \to \mathbb{R}$ une sous-variété constituée des 1-jets de la fonction f en tous les points de V.

Donc le 1-graphe d'une fonction de n variables est une surface de dimension n dans un espace de dimension $2n + 1$.

Théorème. *La forme de contact standard sur la variété des* 1-*jets des fonctions de n variables s'annule en tous les plans tangents aux* 1-*graphes des fonctions. L'adhérence de la réunion des plans tangents à tous les* 1-*graphes des fonctions coïncide avec le sous-espace nul de cette forme (en chaque point de l'espace des jets).*

◀ La première partie découle de la définition de la différentielle totale $du = pdx$, la deuxième, de l'existence d'une fonction possédant des dérivées partielles quelconques en un point. ▶

Le théorème nous apprend qu'un champ de plans défini par une forme de contact standard dans l'espace des 1-jets ne dépend pas du système de coordonnées qui a servi à définir cette forme. Ceci nous conduit à la définition suivante.

Définition. On appelle *structure de contact standard* sur la variété des 1-jets des fonctions sur V le champ d'hyperplans, réunion des plans tangents aux graphes des fonctions sur V.

Exercice. Les groupes de difféomorphismes des espaces V et $\mathbb{R}$ opèrent naturellement sur la variété des 1-jets des fonctions $J^1(V, \mathbb{R})$. Montrer que *la structure de contact standard de l'espace des* 1-*jets est invariante par ces difféomorphismes.*

En outre la 1-forme de contact standard α est invariante par le groupe de difféomorphismes de l'espace V. Elle est multipliée par une fonction non nulle par les difféomorphismes de la droite $\mathbb{R}$.

C. Géométrie sur une hypersurface dans une variété de contact. Revenons à la variété de contact générale M^{2n+1}. Soit E^{2n} une hypersurface différentiable dans M^{2n+1} (fig. 54).

Définition. Une surface E^{2n} d'une variété de contact M^{2n+1} est *non caractéristique* si le plan tangent et le plan de contact en chacun de ses points x sont transversaux (c'est-à-dire leur somme est l'espace tangent à M^{2n+1} ou encore leur intersection est un espace de dimension $2n - 1$).

Définition. L'intersection du plan tangent à une hypersurface non caractéristique et du plan de contact en un point de l'hypersurface dans une variété de contact s'appelle *plan caractéristique* en ce point:

$$P_x = T_xE^{2n} \cap \Pi_x.$$

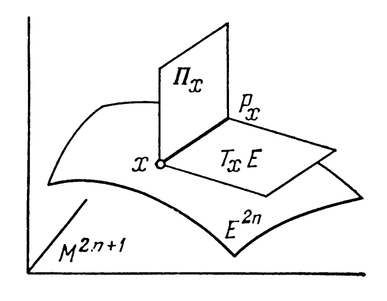


Fig. 54

Donc, les plans caractéristiques sur une hypersurface E^{2n} de M^{2n+1} constituent un champ de plans de dimension $2n - 1$ sur une variété de dimension $2n$. C'est le champ de plans traces des plans de contact dans les espaces tangents à l'hypersurface E^{2n}.

La structure de contact définit sur chacun de ces plans de dimension $2n - 1$ une droite privilégiée appelée *direction caractéristique*.

D. Orthocompléments gauches.

Pour définir la direction caractéristique on rappelle que dans un espace vectoriel de forme bilinéaire non dégénérée sont définis des compléments orthogonaux (deux vecteurs sont orthogonaux si la forme prend la valeur zéro sur ces vecteurs).

Exemple 1. Soit (L, ω) un espace euclidien, c'est-à-dire L est un espace vectoriel, et ω le produit scalaire. A tout vecteur ξ est associée la 1-forme $\omega(\xi, \cdot)$, produit scalaire par le vecteur ξ. Cette 1-forme prend la valeur $\omega(\xi, \eta)$ sur le vecteur η.

Par exemple, grad f est le vecteur associé à la 1-forme df.

Toute droite de L est associée à son orthocomplément (= plan des zéros de la 1-forme associée au vecteur de cette droite). Tout plan de codimension 1 de L est orthocomplément d'une droite (fig. 55).

On remarquera que la multiplication de ω par un nombre non nul préserve l'orthogonalité des vecteurs. Donc la correspondance

entre les droites et leurs orthocompléments est respectée par la multiplication du produit scalaire ω par un nombre non nul.

E x e m p l e 2. Soit (L, ω) un espace symplectique, c'est-à-dire L est un espace vectoriel et ω un produit scalaire gauche (une forme bilinéaire symétrique gauche non dégénérée).

A tout vecteur ξ est associée la 1-forme $\omega(\xi, \cdot)$, produit scalaire par le vecteur ξ. Cette 1-forme prend la valeur $\omega(\xi, \eta)$ sur le vecteur η (fig. 56).

E x e m p l e: le champ hamiltonien de hamiltonien H est un champ associé à la 1-forme dH.

Toute droite de L est associée à son *orthocomplément gauche* (= plan des zéros de la 1-forme associée au vecteur de cette droite).

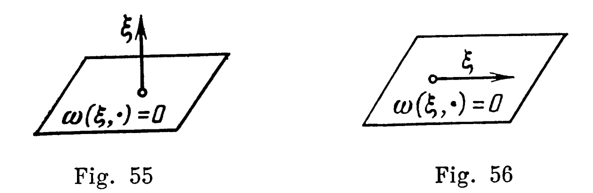


Fig. 55 Fig. 56

Tout plan de codimension 1 de L est l'orthocomplément gauche d'une droite et d'une seule.

On remarquera que l'orthogonalité est préservée par la multiplication de ω par un nombre non nul. Donc la correspondance entre les droites et leurs orthocompléments gauches n'est pas modifiée par la multiplication de la structure symplectique ω par un nombre non nul.

E x e r c i c e 1. Montrer que *toute droite est située dans son orthocomplément gauche.*

S o l u t i o n. $\omega(\xi, \xi) = -\omega(\xi, \xi) = 0$.

E x e r c i c e 2. Montrer que *si tous les vecteurs d'un sous-espace d'un espace symplectique de dimension $2n$ sont orthogonaux gauches, ce sous-espace est de dimension $\leqslant n$.*

S o l u t i o n. L'orthocomplément gauche d'un sous-espace de dimension k est de dimension $2n - k$ (en effet, prenons une base $(e_1, \ldots, e_k)$; alors les équations $\omega(e_1, \xi) = 0, \ldots, \omega(e_k, \xi) = 0$ forment k équations indépendantes en ξ, puisque leur dépendance entraînerait celle des vecteurs e_1 vu la non-dégénérescence de la forme ω).

Si $k > n$ l'orthocomplément est de dimension $< n$ et l'espace ne peut être orthogonal gauche à lui-même.

R e m a r q u e. Dans un espace symplectique de dimension $2n$ il existe des sous-espaces orthogonaux gauches à eux-mêmes. On les appelle *sous-espaces lagrangiens.*

E. Caractéristiques sur une hypersurface d'un espace de contact.

Revenons à la géométrie de l'hypersurface non caractéristique d'une variété de contact M^{2n+1}.

En tout point x de cette hypersurface E^{2n} nous avons défini un plan caractéristique P_x de dimension $2n - 1$, intersection du plan de contact Π_x et du plan T_xE^{2n} tangent en x à E^{2n} (fig. 57).

D é f i n i t i o n. On appelle *direction caractéristique* en un point non caractéristique x d'une hypersurface d'un espace de contact l'orthocomplément gauche du plan caractéristique P_x relativement au plan de contact Π_x (le produit scalaire gauche est défini comme $d\alpha|_{\alpha=0}$) :

l_x = orthocomplément gauche de $P_x^{2n-1} = T_xE^{2n} \cap \Pi_x^{2n}$ relativement à Π_x^{2n}.

A noter que la direction caractéristique en x est contenue dans le plan caractéristique P_x (voir point D); pour $n = 1$ la direction caractéristique coïncide tout simplement avec le plan caractéristique.

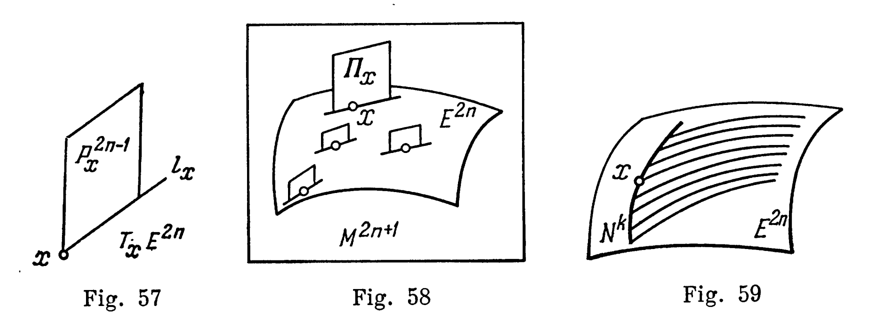


Fig. 57 Fig. 58 Fig. 59

Dans le cas général les directions caractéristiques de tous les points d'une hypersurface non caractéristique différentiable d'un espace de contact forment un champ de directions différentiable sur cette hypersurface; la direction caractéristique en un point appartient aussi à l'hyperplan de contact en ce point (fig. 58).

D é f i n i t i o n. Les courbes intégrales d'un champ de directions caractéristiques défini sur une hypersurface non caractéristique E^{2n} d'une variété de contact sont dites *caractéristiques de* E^{2n}.

Soit $N^k \subset E^{2n}$ une variété intégrale de dimension k du champ des plans de contact.

D é f i n i t i o n. Un point $x \in N^k$ est *non caractéristique* si le plan tangent à N^k en ce point ne contient pas la direction caractéristique (fig. 59).

E x e r c i c e. Démontrer que si x est point non caractéristique d'une variété N^k, alors $k \leqslant n - 1$; citer un exemple de surfaces intégrales de dimension n du champ des plans de contact de M^{2n+1}, contenues dans E^{2n}.

Les caractéristiques passant par les points d'une variété intégrale non caractéristique N^k forment (localement) une sous-variété *intégrale* de dimension $k + 1$ du champ des plans de contact, con-

tenue dans E^{2n}. Pour le prouver on aura besoin d'un lemme général simple sur l'invariance d'un champ de plans par un groupe à un paramètre de difféomorphismes.

F. Digression: condition d'invariance d'un champ de plans. Soit a une 1-forme différentielle partout non nulle. Une telle forme définit un champ d'hyperplans. Soit v un champ de vecteurs partout non nul. Un tel champ définit un champ de directions (c'est-à-dire un champ de droites).

On suppose que la direction du champ v en chaque point appartient à l'hyperplan des zéros de la forme a en ce point

$$a(v) \equiv 0.$$

L e m m e. *Une condition nécessaire et suffisante pour que le champ des zéros de la forme a soit invariant par le flot du champ v est que $da(v, \xi) = 0$ pour tous les points ξ du plan $a(\xi) = 0$ du champ.*

◀ L'assertion du lemme est locale et invariante par les difféomorphismes, donc il suffit de la prouver pour le champ standard $v = \partial/\partial x_1$ dans un espace euclidien rapporté aux coordonnées $(x_1, \ldots, x_n)$ (en vertu du théorème de redressement d'un champ de vecteurs). Soit $a = a_1 dx_1 + \ldots + a_n dx_n$. Par hypothèse, $a_1 \equiv 0$ (puisque $a(v) \equiv 0$).

E x e r c i c e. Montrer que la dérivée extérieure de la forme a prend sur le couple (v, ξ) (où $v = \partial/\partial x_1$, $a(v) \equiv 0$ et ξ, vecteur quelconque) la même valeur que la dérivée partielle $\partial a/\partial x_1$ sur le vecteur ξ.

S o l u t i o n. $\left(da = \sum \frac{\partial a_i}{\partial x_j} dx_j \wedge dx_i\right) \Rightarrow \left(da(v, \xi) = \sum \frac{\partial a_i}{\partial x_1} \xi_i - \sum \frac{\partial a_1}{\partial x_j} \xi_j\right)$. Or $a_1 \equiv 0$.

[On obtient une autre solution plus suggestive en appliquant la formule de Stokes au parallélogramme de côtés v, ξ.]

La condition d'invariance du champ de zéros de la forme a par les translations le long de l'axe x_1 se traduit par la nullité de la dérivée partielle $\partial a/\partial x_1$ sur les plans $a = 0$ du champ.

Sachant que $\partial a/\partial x_1(\xi) = da(v, \xi)$ on voit que la condition d'invariance s'écrit

$$(a(\xi) = 0) \Rightarrow (da(v, \xi) = 0). ▶$$

G. Problème de Cauchy pour champ de directions caractéristiques. Revenons à l'hypersurface non caractéristique E^{2n} de la variété de contact M^{2n+1}. Soit $N^k \subset E^{2n}$ une sous-variété intégrale du champ des plans de contact.

D é f i n i t i o n. Le *problème de Cauchy* pour l'hypersurface E^{2n} de M^{2n+1} avec la variété initiale N^{n-1} consiste à trouver une variété intégrale Y^n du champ des plans de contact, appartenant à E^{2n} et contenant la variété initiale N^{n-1} (fig. 60).

Théorème. *Soit x un point non caractéristique* *) *d'une variété initiale N^{n-1}. Il existe alors un voisinage U du point x tel que la solution du problème de Cauchy pour $E^{2n} \cap U$ avec la condition initiale $N^{n-1} \cap U$ existe et est unique localement (c'est-à-dire deux solutions quelconques vérifiant la même condition initiale coïncident dans un voisinage du point x).*

La variété Y^n est constituée de caractéristiques passant par les points de la variété initiale N^{n-1}.

◀ La famille de caractéristiques menées par les points de la variété initiale forment au voisinage du point x une sous-variété diffé-

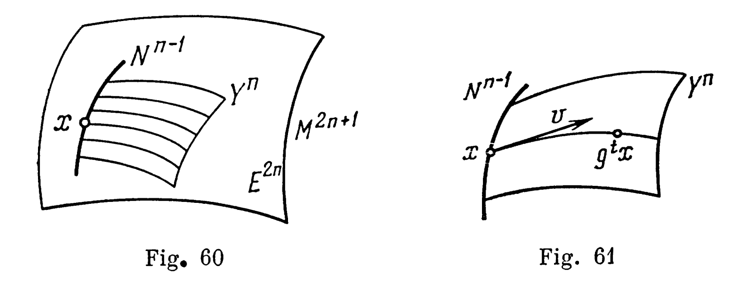


Fig. 60 Fig. 61

rentiable Y^n de dimension n dans E^{2n}. Montrons que cette variété est intégrale pour le champ des hyperplans de contact.

Soit une 1-forme de contact α définissant le champ d'hyperplans de contact au voisinage du point x. Désignons par a la restriction de la forme α au voisinage du point x dans E^{2n}.

La forme a ne s'annule pas puisque l'hypersurface E^{2n} est non caractéristique (voir point C). Cette forme définit sur E^{2n} un champ d'hyperplans tangents (qui sont les traces des hyperplans de contact sur E^{2n}). Le champ de directions caractéristiques sur E^{2n} est contenu dans le champ de plans des zéros de la forme a (voir point E).

Considérons au voisinage du point x de E^{2n} un champ de vecteurs dont la direction est caractéristique en chaque point. Appelons $\{g^t\}$ son groupe local à un paramètre de difféomorphismes sur E^{2n} (g^t est défini au voisinage du point x pour les valeurs de t proches de zéro) (fig. 61).

Tout point de la variété Y^n est image d'un point de la variété initiale par un difféomorphisme convenable g^t.

Lemme A. *La forme a est nulle sur tout plan tangent à Y aux points de la variété initiale.*

Lemme B. *Les plans des zéros de la forme a se correspondent par les difféomorphismes g^t.*

A. ◀ La forme a est nulle sur les vecteurs tangents à N^{n-1}, puisque la variété N est intégrale. La forme a est nulle sur le vecteur v,

*) Le point non caractéristique a été défini au point E.

puisque la direction caractéristique est contenue dans un hyperplan de contact. Donc la forme a est nulle sur la somme $T_x N + \mathbb{R}v$. ▶

B. ◀ Soit un vecteur ξ sur lequel s'annule la forme a (le vecteur ξ n'a pas obligatoirement son origine en un point de la variété initiale). Calculons la valeur prise par da sur le couple (v, ξ) de vecteurs tangents à E^{2n} en un même point. Par définition, la forme $a = \alpha\,|_E$ est la valeur de la dérivée de la forme de contact $d\alpha\,(v, \xi)$. Par définition de la direction caractéristique, cette valeur est nulle sur tout vecteur ξ dans un plan de contact. Le lemme du point F nous dit que $a = 0$ est invariant par $\{g^t\}$. ▶

Des lemmes A et B il suit que la forme a s'annule sur tous les vecteurs tangents à Y^n, donc la variété Y^n est intégrale.

Nous avons donc construit une sous-variété intégrale $Y^n \subset E^{2n}$ d'un champ d'hyperplans de contact, passant par la variété initiale N^{n-1}.

H. Démonstration de l'unicité.

L e m m e. *Tout plan tangent à une variété intégrale de dimension n du champ des plans de contact contenus dans E^{2n} renferme une direction caractéristique.*

◀ La restriction de la 1-forme de contact α à l'une quelconque de ses variétés intégrales est nulle. La restriction de la 2-forme $d\alpha$ est égale à la dérivée de la restriction de la forme α, c'est-à-dire à 0. Donc les vecteurs tangents à la variété intégrale sont deux à deux orthogonaux gauches (au sens du produit scalaire $d\alpha\,|_{\alpha=0}$).

Chaque vecteur de direction caractéristique sur E^{2n} est orthogonal gauche à tous les vecteurs de E^{2n} contenus dans le plan $\alpha = 0$. Supposons que ce vecteur n'appartient pas au plan tangent en son origine à la variété intégrale de dimension n du champ de plans de contact, contenue dans E^{2n}. Le sous-espace engendré par ce vecteur et ce plan tangent est alors de dimension $n + 1$. Or tous les vecteurs du sous-espace construit sont deux à deux orthogonaux gauches. Donc en vertu de l'exercice 2 du point D, la dimension de ce sous-espace est $\leqslant n$. ▶

Le lemme nous dit que toute variété intégrale de dimension n du champ des plans de contact, contenue dans E^{2n}, renferme avec chacun de ses points un intervalle de caractéristique passant par ce point. Ceci entraîne l'unicité de la variété intégrale renfermant la variété initiale considérée. ▶

I. Application à l'équation non linéaire aux dérivées partielles du premier ordre.

Considérons maintenant l'équation non linéaire aux dérivées partielles du premier ordre en la fonction $u\colon V^n \to \mathbb{R}$ comme une hypersurface E^{2n} de la variété des 1-jets $M^{2n+1} = J^1(V^n, \mathbb{R})$, munie de la structure de contact standard.

Soient $(x_1, \ldots, x_n)$ des coordonnées locales sur V^n, u une coordonnée locale sur $\mathbb{R}$; $(x_1, \ldots, x_n, u; p_1, \ldots, p_n)$ les coordonnées locales correspondantes sur l'espace des 1-jets. L'équation différentielle s'écrit alors (localement)

$$\Phi(x, u, p) = 0. \tag{1}$$

L'intégration de cette équation se ramène à la détermination dans E^{2n} de surfaces intégrales du champ des plans de contact, 1-graphes de fonctions (cf. point B).

Les théorèmes généraux ramènent l'intégration de cette équation à la construction de caractéristiques sur E^{2n}. Pour construire ces caractéristiques il faut trouver les courbes intégrales d'un champ de directions sur E^{2n} (c'est-à-dire résoudre un système de $2n - 1$ équations différentielles ordinaires).

Théorème. *Les solutions de l'équation* (1) *sont des fonctions dont les 1-graphes sont les réunions des caractéristiques de E^{2n}.*

◀Voir points B, G, I.▶

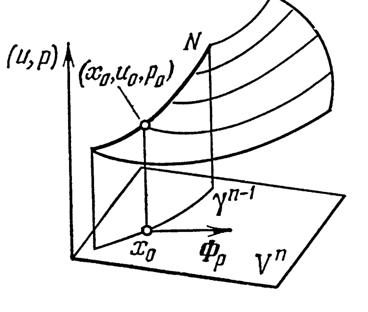


Fig. 62

J. Problème de Cauchy pour l'équation non linéaire aux dérivées partielles du premier ordre.

Soient $\gamma^{n-1} \subset V^n$ une sous-variété de dimension $n - 1$ d'une variété V^n; $\varphi: \gamma^{n-1} \to \mathbb{R}$ une fonction différentiable; $E^{2n} \subset J^1(V^n, \mathbb{R})$ une surface non caractéristique différentiable définie par l'équation (1).

Définition. Le *problème de Cauchy* pour l'équation (1) consiste à trouver la solution $u: V \to \mathbb{R}$ qui coïncide avec φ sur γ.

Définition. On appelle *variété initiale N* construite d'après la condition (γ, φ) l'ensemble composé de tous les 1-jets des fonctions sur V^n vérifiant les conditions suivantes (fig. 62):

1) Le point d'application d'un jet appartient à γ^{n-1}.

2) La fonction prend la valeur φ en ce point.

3) La différentielle totale de la fonction en ce point est telle que sa restriction au plan tangent à γ^{n-1} est égale à la différentielle totale de la condition initiale φ.

4) Le jet appartient à E^{2n}.

Définition. Un point d'une variété initiale est dit *non caractéristique* pour l'équation (1) si la projection de la direction caractéristique en ce point sur V est transversale à γ. (Cette définition est différente de celle donnée au point E.)

Remarque. On dit que les dérivées de la fonction inconnue suivant $n - 1$ directions le long de γ sont définies par la condition

initiale et que la dérivée par rapport à la dernière direction (qui est transversale à γ) est déterminée à partir de l'équation (1).

E x e m p l e. Supposons que γ est définie par l'équation $x_1 = 0$ dans un espace muni des coordonnées (x_1, x'). Alors N est donné par les conditions

$$x_1 = 0, \quad u = \varphi(x'), \quad p' = \partial\varphi/\partial x';$$

p_1 se déduit à partir de l'équation $\Phi(x, u, p) = 0$.

T h é o r è m e. *Supposons que le point* (x_0, u_0, p_0) *de l'espace des 1-jets est un point non caractéristique de la variété initiale* N. *Alors l'équation* (1) *avec la condition initiale* N *admet en un voisinage* U *du point initial* x_0 *une solution localement unique* (*l'unicité est comprise au sens que deux solutions vérifiant la condition initiale* $u|_{U\cap\gamma} = \varphi|_{U\cap\gamma}$, $u(x_0) = u_0$, $du(x_0) = p_0$ *coïncident en un voisinage du point* x_0).

◀ Ceci découle du théorème du point G qui nous donne une méthode de construction de cette solution. ▶

K. Formules explicites.

E x e r c i c e. Expliciter l'équation différentielle des caractéristiques pour l'équation $\Phi(x, u, p) = 0$.

R é p o n s e.

$$\dot{x} = \Phi_p,$$
$$\dot{p} = -\Phi_x - p\Phi_u,$$
$$\dot{u} = p\Phi_p.$$

S o l u t i o n. Considérons un vecteur tangent à la variété $\Phi = 0$. Si (X, U, P) sont les composantes de ce vecteur, on a

$$\Phi_x X + \Phi_u U + \Phi_p P = 0.$$

Ce vecteur appartient au plan de contact en son origine si la forme $du - p\,dx$ est nulle, c'est-à-dire $U = pX$.

Donc, une condition nécessaire et suffisante pour que le vecteur (X, pX, P) appartienne à l'intersection du plan de contact et du plan tangent en son origine à la variété $\Phi = 0$ est

$$|\,\Phi_x + \Phi_u p\,|\, X + \Phi_p P = 0. \tag{2}$$

Le vecteur caractéristique $(\dot{x}, \dot{u} = p\dot{x}, \dot{p})$ est déterminé à partir de la nullité du produit scalaire gauche de ce vecteur par tous les vecteurs (X, pX, P) satisfaisant l'égalité précédente.

Or le produit scalaire gauche est égal à la valeur prise par $d\alpha = dx \wedge dp$ sur le couple de vecteurs $(\dot{x}, \dot{u} = p\dot{x}, \dot{p})$, $(X, U = pX, P)$, c'est-à-dire à $\dot{x}P - \dot{p}X$.

Donc l'équation (2) en X et P doit être équivalente à l'équation

$$\dot{x}P - \dot{p}X = 0. \tag{3}$$

Par suite, les coefficients en X et P des équations (2) et (3) sont proportionnels. D'où la réponse, puisque $\dot{u} = p\dot{x}$.

L. Conditions pour les points non caractéristiques.

E x e r c i c e. Expliciter les conditions imposées à γ, φ et Φ pour qu'un point (x_0, u_0, p_0) soit non caractéristique pour l'équation $\Phi(x, u, p) = 0$ avec la condition initiale φ sur γ.

R é p o n s e. $\Phi_p(x_0, u_0, p_0)$ ne doit pas être tangent à γ en x_0 (cf. fig. 62, page 83).

S o l u t i o n. Projetons le plan tangent en (x_0, u_0, p_0) à la surface formée par les caractéristiques passant par la variété initiale sur l' x-espace. Si cette surface est 1-graphe de la solution et le point (x_0, u_0, p_0) non caractéristique, alors le plan tangent à cette surface en (x_0, u_0, p_0) est engendré par le plan tangent à N en (x_0, u_0, p_0) et une direction caractéristique et se projette isomorphement. Par suite, la x-composante du vecteur caractéristique doit être transversale à γ en x_0. Or cette composante n'est autre que Φ_p (cf. point K).

Inversement, supposons que Φ_p n'est pas tangent à γ au point x_0. Alors

1) *L hypersurface $\Phi = 0$ est différentiable au voisinage du point (x_0, u_0, p_0).* En effet, $\Phi_p \neq 0$ et, par suite, $d\Phi|_{(x_0, u_0, p_0)} \neq 0$.

2) *L'équation $\Phi = 0$ n'est pas caractéristique au point (x_0, u_0, p_0).* En effet, le vecteur $(0, 0, \Phi_p)$ est situé dans un plan de contact et n'est pas tangent à la surface $\Phi = 0$ au point (x_0, u_0, p_0), puisque $\Phi_p(x_0, u_0, p_0) \neq 0$.

3) *La variété initiale N est différentiable au voisinage du point (x_0, u_0, p_0).* En effet, choisissons les coordonnées $(x_1, \ldots, x_n) = (x_1, x')$ telles que l'équation locale de γ prenne la forme $x_1 = 0$. La condition de résolubilité de l'équation en $p_1(x')$

$$\Phi(0, x', \varphi(x'), p_1, \partial\varphi/\partial x') = 0$$

s'écrit alors

$$\partial\Phi/\partial p_1|_{(x_0, u_0, p_0)} \neq 0;$$

la condition de non-tangence du vecteur Φ_p à γ est de la même forme.

4) *Le point (x_0, u_0, p_0) de la variété initiale est non caractéristique.*

En effet, si le vecteur caractéristique était tangent à la variété initiale N, sa projection Φ_p serait tangente à la projection γ de la variété N sur l'x-espace.

5) *Les caractéristiques coupant la variété initiale au voisinage du point (x_0, u_0, p_0) forment dans ce voisinage une variété différentiable qui se projette difféomorphement sur l'x-espace (et, par conséquent, est 1-graphe d'une fonction).*

En effet, l'image du plan tangent à cette variété au point (x_0, u_0, p_0) par projection sur l'x-espace contient le plan tangent en x_0 à γ et le vecteur qui lui est transversal. Donc la dérivée de l'application étudiée au point (x_0, u_0, p_0) est un isomorphisme, quant à la projection, elle est difféomorphisme local (d'après le théorème de la fonction inverse).

Ainsi, les conditions 1)-5) sont réalisées au point (x_0, u_0, p_0), si Φ_p n'y est pas tangente à γ.

M Equation de Hamilton-Jacobi.

D é f i n i t i o n. On appelle *équation de Hamilton-Jacobi* l'équation

$$H(x, u_x) = 0. \qquad (1)$$

Cette équation se distingue de l'équation générale à dérivées partielles de premier ordre par le fait que la valeur de la fonction inconnue n'y figure pas explicitement.

E x e m p l e. Soient γ une hypersurface différentiable dans l'espace euclidien $\mathbb{R}^n$, $u(x)$ la distance du point x à γ (fig. 63). Alors la fonction u vérifie (en ses points de différentiabilité) l'équation de

Hamilton-Jacobi

$$\left(\frac{\partial u}{\partial x_1}\right)^2+\ldots+\left(\frac{\partial u}{\partial x_n}\right)^2=1. \qquad (2)$$

En effet, le module du gradient de cette fonction est égal à celui de la dérivée de la distance à γ par rapport à la normale à γ, c'est-à-dire à 1.

En général la fonction u peut ne pas être différentiable. Par exemple, soit γ une ellipse dans le plan. Les singularités de u forment alors un segment à l'intérieur de l'ellipse (fig. 64).

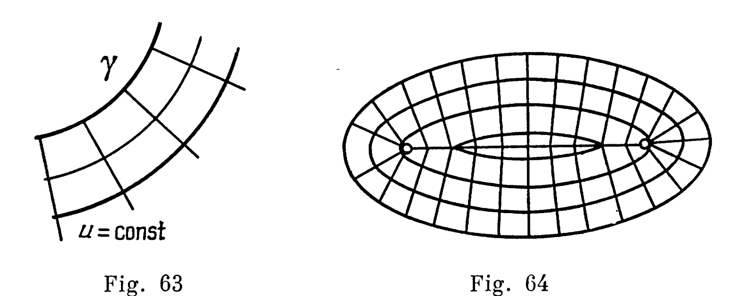


Fig. 63 Fig. 64

E x e r c i c e. Montrer que toute solution de l'équation de Hamilton-Jacobi (2) est localement la somme de la distance à une hypersurface et d'une constante.

Lorsqu'on étudie l'équation de Hamilton-Jacobi on a intérêt à considérer le *fibré cotangent* T^*V^n au lieu de la variété des 1-jets $J^1(V^n, \mathbb{R})$ des fonctions sur V^n. En mécanique, l'espace T^*V^n est appelé *espace des phases* pour l'espace de configuration V^n. On appelle *vecteur cotangent* à V^n en x une fonction homogène linéaire définie sur l'espace tangent à V^n en x. Tous les vecteurs cotangents à V^n en x forment un espace vectoriel appelé *espace cotangent* à V^n en x et noté $T_x^*V^n$. Les vecteurs cotangents à V^n en tous ses points forment une variété différentiable de dimension $2n$ appelée *espace du fibré cotangent* à V^n (ou, tout simplement, fibré cotangent) et notée T^*V^n.

Soient $(x_1, \ldots, x_n)$ des coordonnées locales sur V^n. Un vecteur cotangent à V^n en x est défini par une collection de n nombres $(p_1, \ldots, p_n)$. Plus exactement, à la collection des nombres $\{p_k\}$ est associée la 1-forme $p_1dx_1 + \ldots + p_ndx_n$ sur l'espace tangent à V^n en x. La collection de $2n$ nombres $(p_1, \ldots, p_n; x_1, \ldots, x_n)$ forme un système de coordonnées locales dans le fibré cotangent à V^n.

Il existe une projection naturelle π de $J^1(V^n, \mathbb{R})$, espace des 1-jets des fonctions sur le fibré cotangent à V^n

$$\pi: J^1(V^n, \mathbb{R}) \to T^*V^n.$$

L'application π consiste à « oublier les valeurs de la fonction »; en coordonnées elle s'écrit:

$$(x_1, \ldots, x_n; u; p_1, \ldots, p_n) \longmapsto (p_1, \ldots, p_n; x_1, \ldots, x_n).$$

D é f i n i t i o n. On appelle *caractéristiques de l'équation de Hamilton-Jacobi* (1) les projections des caractéristiques de l'équation à dérivées partielles de premier ordre (1) sur le fibré cotangent.

E x e r c i c e. Trouver l'équation différentielle des caractéristiques de l'équation de Hamilton-Jacobi (1).

R é p o n s e. $\dot{x} = H_p$, $\dot{p} = -H_x$.

R e m a r q u e. Ce système d'équations différentielles s'appelle *système des équations canoniques de Hamilton.* Le champ de vecteurs correspondant est défini aussi bien sur la surface $H = 0$ que sur l'espace des phases tout entier.

E x e r c i c e. Trouver les caractéristiques de l'équation de Hamilton-Jacobi (2).

R é p o n s e. $x = 2at + b$, $p = a$ (a et b sont des vecteurs constants, $a^2 = 1$).

Donc, les projections des caractéristiques sur V^n sont des droites.

En optique géométrique l'équation de Hamilton-Jacobi (2) est appelée *équation de l'iconal*, les projections des caractéristiques sur V^n,

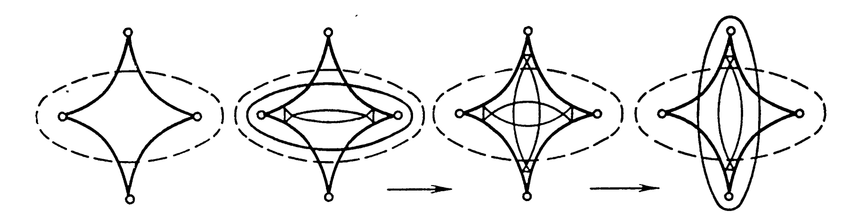


Fig. 65

rayons, la fonction *u longueur optique du chemin*, les surfaces de niveau de u, *fronts*. Par ailleurs les *caustiques* sont essentielles en géométrie optique. Considérons, par exemple, une paroi éclairée par des rayons réfléchis par une surface concave (l'intérieur d'une tasse). Sur la paroi on voit se dessiner des lignes plus brillantes avec des points singuliers: ce sont les caustiques.

Les caustiques se définissent comme suit. Considérons le problème de Cauchy pour une équation aux dérivées partielles de premier ordre. Même si les caractéristiques correspondantes sont indéfiniment prolongeables et ne se coupent pas de sorte à former une variété intégrale globale, la projection de cette variété dans V^n n'est généralement pas un difféomorphisme.

L'ensemble des valeurs critiques de la projection de la variété intégrale dans V^n s'appelle *caustique*.

Dans le cas particulier de l'équation de Hamilton-Jacobi (2) avec la condition initiale $u = 0$ sur γ la caustique est le lieu géométrique des foyers ou des centres de courbure de l'hypersurface γ.

Exercice 1. Tracer le lieu géométrique des centres de courbure de l'ellipse.

Exercice 2. Porter un segment de longueur t sur toute normale intérieure à une ellipse. Représenter la courbe obtenue et étudier ses variations lorsque t croît.

Réponses. Voir figure 65.

§ 9. Théorème de Frobénius

Un champ de directions sur le plan définit toujours une famille de courbes intégrales et est localement rectifiable (il se ramène à un champ de plans parallèles par un difféomorphisme). Ce n'est plus le cas à partir des espaces à trois dimensions : un champ de plans dans $\mathbb{R}^3$ peut ne pas présenter de surfaces intégrales.

Dans ce paragraphe on étudie les conditions de rectification locale d'un champ d'hyperplans, c'est-à-dire les conditions sous lesquelles un champ est champ de tangentes à une famille d'hypersurfaces différentiables.

A. Champ d'hyperplans complètement intégrable.

Soit M^n une variété différentiable sur laquelle est défini un champ d'hyperplans tangents.

Un tel champ définit au voisinage de tout point une 1-forme différentielle α non nulle déterminée au produit près par une fonction partout non nulle.

Définition. Un champ d'hyperplans est *complètement intégrable* si la forme $d\alpha$ est identiquement nulle dans tous les plans de ce champ.

Remarque. La propriété d'intégrabilité complète d'un champ ne dépend pas du choix de la forme α qui a servi à le définir localement, puisque si l'on multiplie la forme α par une fonction non nulle la forme $d\alpha\ |_{\alpha=0}$ est multipliée par cette fonction (voir § 8, point A).

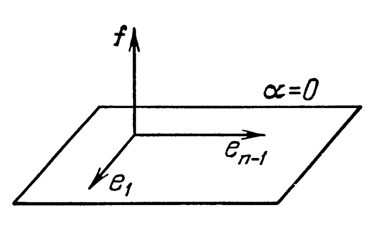


Fig. 66

Proposition. *Une condition nécessaire et suffisante pour qu'un champ d'hyperplans $\alpha = 0$ soit complètement intégrable est que*

$$\alpha \wedge d\alpha \equiv 0.$$

◀ Prenons dans l'espace tangent à M^n u point considéré une base (fig. 66) constituée d'un vecteur $n - 1$ « horizontal » ($e_1, \ldots$ $\ldots, e_{n-1}$) d'un plan $\alpha = 0$ et d'un vecteur « vertical » f. La 3-forme $\alpha \wedge d\alpha$ est nulle sur trois vecteurs horizontaux puisque $\alpha = 0$.

D'autre part, $(d\alpha \wedge \alpha)(e_i, e_j, f) = 0$ comme somme dans laquelle chaque terme est multiplié soit par $\alpha(e_i)$, soit par $d\alpha(e_i, e_j)$ qui sont nulles.

Inversement, si $\alpha \wedge d\alpha = 0$, alors $d\alpha(e_i, e_j) = 0$. En effet, l'unique terme de $(\alpha \wedge d\alpha)(e_i, e_j, f)$ à ne pas contenir $\alpha(e_i)$ ou $\alpha(e_j)$ est $d\alpha(e_i, e_j) \cdot \alpha(f)$; or $\alpha(f) \neq 0$. ▶

R e m a r q u e. La condition $\alpha \wedge d\alpha \equiv 0$ s'appelle *condition d'intégrabilité de Frobénius.* De la proposition démontrée il suit que c'est une condition portant sur le champ de plans qui est réalisée ou non simultanément pour toutes les formes définissant le champ.

B. Existence de variétés intégrales.

T h é o r è m e. *Pour qu'un champ d'hyperplans $\alpha = 0$ soit champ de tangentes à une famille d'hypersurfaces il faut et il suffit que ce champ vérifie la condition d'intégrabilité de Frobénius $\alpha \wedge d\alpha \equiv 0$.*

◀ Sur les surfaces de la famille $\alpha = 0$, donc $d\alpha = 0$. Inversement, supposons que $d\alpha = 0$ dans les plans $\alpha = 0$. Nous allons construire une famille de surfaces intégrales au voisinage d'un point x de la manière suivante. Soit v un champ de vecteurs tel que $\alpha(v) \equiv 0$ (c'est-à-dire qu'en chaque point le vecteur du champ est contenu dans un plan du champ). Soit Γ^k une sous-variété intégrale du champ de plans (fig. 67) et supposons que $v(x)$ n'est pas contenu dans le plan tangent à Γ^k en x.

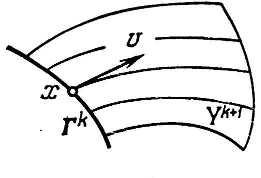


Fig. 67

L e m m e. *Les trajectoires du champ v passant par les points de la variété intégrale Γ^k forment au voisinage du point x une variété intégrale différentiable Y^{k+1} du champ de plans complètement intégrable $\alpha = 0$.*

◀ Désignons par $\{g^t\}$ le flot local du champ v. Alors

a) *Les plans du champ $\alpha = 0$ se correspondent par les difféomorphismes g^t.*

En effet, $d\alpha(\xi) = 0$ pour tout vecteur ξ d'un plan du champ, donc le champ est invariant par les difféomorphismes g^t d'après le lemme du point F, § 8.

b) *L'espace tangent à Y^{k+1} aux points de la variété initiale Γ^k est contenu dans un plan du champ.*

En effet, le plan tangent en x à la variété intégrale Γ^k et le vecteur v sont situés dans un plan du champ, quant à l'espace tangent à Y^{k+1} en x il est engendré par $v(x)$ et $T_x\Gamma^k$.

De a) et b) il suit que la variété Y^{k+1} est intégrale pour le champ de plans $\alpha = 0$. ▶

On peut construire maintenant les variétés intégrales de dimension $n - 1$ en accroissant successivement la dimension.

Considérons un système de coordonnées locales $(x_1, \ldots, x_{n-1}; y)$ dans lequel le plan de coordonnées $y = 0$ au point zéro appartient au champ de plans $\alpha = 0$.

La projection le long de l'axe y à partir d'un plan du champ sur le plan de coordonnées $(x_1, \ldots, x_n)$ au voisinage du zéro est un isomorphisme (fig. 68).

Considérons les champs de vecteurs de base $(\partial/\partial x_1, \ldots, \partial/\partial x_{n-1})$ dans le plan de coordonnées. Leurs antécédents dans les plans du

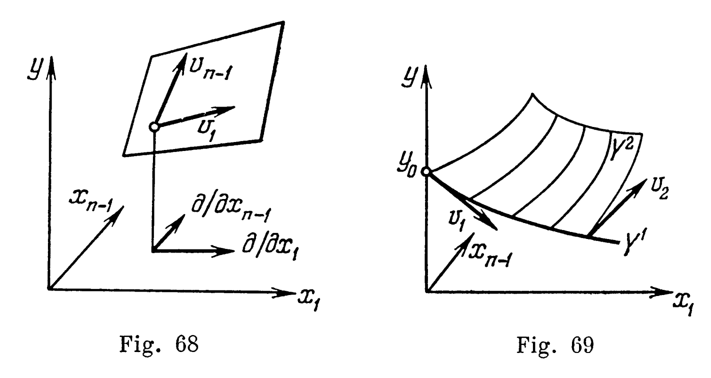


Fig. 68

Fig. 69

champ forment des champs de vecteurs différentiables au voisinage du point O. Désignons ces champs par $(v_1, \ldots, v_{n-1})$.

Prenons le point y_0 de l'axe y pour variété intégrale initiale Γ^0 de dimension zéro (fig. 69).

En appliquant le lemme à Γ^0 et v_1, on obtient une variété intégrale Y^1 de dimension un sur laquelle $x_2 = \ldots = x_{n-1} = 0$. Donc $v_2 \notin T_0Y^1$.

En appliquant le lemme à Y^1 et v_2, on obtient une variété intégrale Y^2 de dimension deux. En poursuivant cette procédure nous arrivons à la variété intégrale Y^k sur laquelle $x_{k+1} = \ldots = x_{n-1} = 0$; en appliquant le lemme à Y^k et à v_{k+1} on obtient une variété intégrale Y^{k+1} sur laquelle $x_{k+2} = \ldots = x_{n-1} = 0$.

Cette procédure s'achève par la construction de la variété cherchée Y^{n-1}. ▶

CHAPITRE 3

STABILITÉ STRUCTURELLE

Un problème qui se pose invariablement quand on se sert d'un modèle mathématique est la validité des résultats relatifs au comportement de ce modèle dans la réalité. En effet si les résultats sont très sensibles aux moindres variations du modèle (par exemple à une petite variation du champ de vecteurs définissant l'équation différentielle), l'application de ces résultats au processus réel présente des dangers, car sa simulation suppose toujours une certaine idéalisation, les paramètres n'étant déterminés qu'approximativement, etc. D'où la nécessité de sélectionner les propriétés du modèle qui sont peu sensibles à une faible variation de ce dernier et, par conséquent, sont susceptibles de caractériser le processus réel.

Une tentative de sélection de ces propriétés a conduit à la notion de stabilité structurelle (A. Andronov et L. Pontriaguine, 1937). Les résultats spectaculaires de la théorie de la stabilité structurelle dans un espace des phases de petite dimension (1 et 2) ont entretenu un espoir qui n'a été dissipé que dans les années 1960 par les travaux de S. Smale qui a prouvé que dans un espace des phases de grande dimension certains systèmes n'admettent pas d'approximation par des systèmes structurellement stables. Ce résultat est pour la théorie qualitative des équations différentielles à peu de chose près ce que le théorème de Liouville relatif à la non-résolubilité des équations différentielles par des quadratures est à la théorie de l'intégration des équations différentielles. Plus exactement, il montre que dans un espace des phases à plusieurs dimensions une classification topologique complète des équations différentielles est impossible même si l'on se limite à des équations génériques et que l'on néglige tous les cas dégénérés.

Ce chapitre passe en revue les notions, méthodes et résultats principaux de la théorie de la stabilité structurelle.

§ 10. Notion de stabilité structurelle

Dans ce paragraphe on définit la stabilité structurelle et on étudie les champs de vecteurs structurellement stables dans un espace de phases à une dimension.

A. Définition naïve de la stabilité structurelle

Soit l'équation différentielle

$$\dot{x} = v(x), \quad x \in M,$$

définie par un champ de vecteurs v sur une variété M. On dira aussi que le champ v définit un *système dynamique* (ou simplement un système). On supposera (généralement) que les solutions de l'équation sont indéfiniment prolongeables: ce sera toujours le cas si M est compacte.

E x e m p l e. L'équation du pendule en présence de frottement (fig. 70):

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = -x_1 - kx_2.$$

Si $k = 0$ toutes les trajectoires sont fermées. Si $k > 0$ elles s'enroulent autour du foyer O. Donc, une légère variation du coefficient

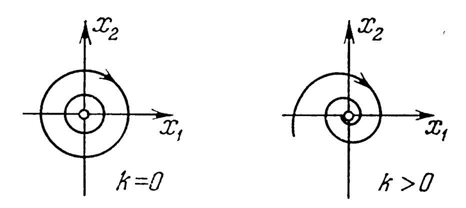


Fig. 70

de frottement modifie ou non le comportement des trajectoires selon que ce coefficient est nul ou positif.

La définition de la stabilité structurelle qui va suivre formalise cette différence: un pendule sans frottement est un système structurellement instable, un pendule avec frottement, structurellement stable.

D é f i n i t i o n. Un système est *structurellement stable* s'il reste équivalent à lui-même par une petite variation du champ de vecteurs.

Pour que cette définition ait un sens il faut dire ce qu'on entend par petite variation du champ et par systèmes équivalents.

B. Equivalence topologique.

La classification la plus fine des équations différentielles repose sur la notion de difféomorphisme. Deux systèmes (M_1, v_1) et (M_2, v_2) sont *difféomorphes* s'il existe un difféomorphisme $h: M_1 \to M_2$ envoyant le champ de vecteurs v_1 dans le champ de vecteurs v_2.

Les systèmes difféomorphes sont absolument identiques du point de vue de la géométrie des variétés différentiables. L'exemple suivant montre que la classification à un difféomorphisme près est trop fine (car elle introduit un trop grand nombre de classes d'équivalence).

E x e m p l e. Soient sur un espace des phases de dimension un les équations $\dot{x} = x$ et $\dot{x} = 2x$.

Dans les deux cas le zéro est l'unique point d'équilibre (répulsif). Cependant ces deux systèmes ne sont pas difféomorphes.

◀ Si un difféomorphisme envoie un point singulier x_1 d'un champ de vecteurs dans un point singulier x_2 d'un autre champ de vecteurs, la dérivée de ce difféomorphisme envoie l'opérateur de la partie linéaire du premier champ en x_1 dans l'opérateur de la partie linéaire du second champ en x_2. Ces deux opérateurs sont donc semblables et possèdent en particulier les mêmes valeurs propres. Par suite, les valeurs propres du linéarisé d'un champ de vecteurs en un point singulier sont des invariants qui varient continûment avec le champ. Ces invariants sont dits *modules*. L'existence des modules conduit au fait que la partition de l'ensemble des champs de vecteurs en classes de champs difféomorphes n'est pas discrète, mais continue (fig. 71).

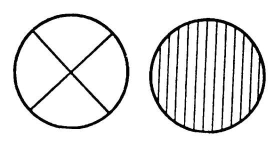


Fig. 71.

En particulier, les champs précédents ne sont pas difféomorphes, puisque $1 \neq 2$. ▶

Pour ne pas distinguer ces deux champs, on introduit une équivalence plus grossière appelée équivalence topologique. A noter que les homéomorphismes (applications biunivoques et bicontinues) n'agissent pas sur les champs de vecteurs. Donc l'équivalence topologique des champs de vecteurs se définit comme suit.

Considérons les flots définis par les champs de vecteurs donnés. Le flot du champ v sur M est constitué des applications $g^t : M \to M$ associant à toute condition initiale x_0 de l'équation $\dot{x} = v(x)$ à l'instant 0 la valeur $g^t x_0$ de cette solution à l'instant t; de toute évidence, $g^{t+s} = g^t g^s$, $g^0 = 1$. Si M est compacte, alors $g^t x$ sont définies pour tous les $t \in \mathbb{R}$ et $x \in M$.

D é f i n i t i o n. Deux systèmes sont *topologiquement équivalents* si existe un homéomorphisme de l'espace des phases d'un système sur l'espace des phases de l'autre système envoyant le flot du premier dans celui du second:

$$hg_1^t x \equiv g_2^t hx.$$

Autrement dit, le diagramme

$$\begin{array}{ccc} M_1 & \xrightarrow{g_1^t} & M_1 \\ h\downarrow & & \downarrow h \\ M_2 & \xrightarrow{g_2^t} & M_2 \end{array}$$

est commutatif.

Par exemple, les systèmes $\dot{x} = x$ et $\dot{x} = 2x$ sont topologiquement équivalents.

R e m a r q u e. La notion d'homéomorphisme et la topologie continue (non différentielle) doivent leur existence à l'application des homéomorphismes pour éliminer les modules.

C. Equivalence orbitale.

Malheureusement la notion d'équivalence topologique des systèmes ne nous délivre pas des modules.

E x e m p l e. Considérons un champ de vecteurs possédant une trajectoire fermée, un cycle limite par exemple. Tout système topologiquement équivalent possèdera un cycle limite de même période. Une petite variation du champ peut entraîner une légère variation de la période. Donc la période du mouvement sur le cycle est un module par rapport à l'équivalence topologique. Pour éliminer ce module on introduit une classification plus grossière encore que celle à un homéomorphisme près.

D é f i n i t i o n. Deux systèmes sont *topologiquement orbitalement équivalents* si existe un homéomorphisme de l'espace des phases d'un système sur l'espace des phases du second associant aux trajectoires orientées du premier des trajectoires orientées du second. Cette définition n'implique pas l'égalité des intervalles de temps sur les portions correspondantes de trajectoires associées.

L'hypothèse de la stabilité structurelle consiste en ce que la partition des systèmes en classes d'équivalence orbitale ne possède plus de module (est discrète) tout au moins si on se limite aux cas génériques et l'on néglige les dégénérescences.

D. Définition finale de la stabilité structurelle.

Soient M une variété différentiable compacte (de la classe C^{r+1}, $r \geqslant 1$), v un champ de vecteurs de la classe C^r (si M est à bord, on suppose que v n'est pas tangent à ce bord).

Le système (M, v) est *structurellement stable* si existe un voisinage du champ v dans C^1 tel que tout champ de vecteurs de ce voisinage définisse un système topologiquement orbitalement équivalent au système initial, l'équivalence étant réalisée par un homéomorphisme proche de l'identité.

E. Cas unidimensionnel.

Soit M un cercle. Un champ de vecteurs sur un cercle est défini par une fonction périodique. Les points singuliers du champ sont associés aux zéros de cette fonction. Un point singulier est *non dégénéré* si la dérivée de la fonction y est non nulle.

T h é o r è m e. *Un champ de vecteurs sur un cercle définit un système structurellement stable si et seulement s'il ne possède que des points singuliers non dégénérés.*

Deux champs de vecteurs à points singuliers non dégénérés sur un cercle sont topologiquement orbitalement équivalents si et seulement s'ils possèdent le même nombre de points singuliers.

Les champs de vecteurs structurellement stables forment dans l'espace de tous les champs de vecteurs sur le cercle un ensemble ouvert partout dense.

◀ Supposons que tous les points singuliers d'un champ sont non dégénérés. Ils sont donc en nombre fini et sont alternativement stables ou instables. Toute solution non constante de l'équation $\dot{x} = v(x)$ tend vers une position d'équilibre stable pour $t \to +\infty$, et instable pour $t \to -\infty$. D'où suivent toutes les propositions du théorème à l'exception de celle qui consiste à prouver que tous les points singuliers du champ peuvent être rendus non dégénérés par une déformation aussi petite que l'on veut du champ.

Il est commode de prouver cette proposition à l'aide du lemme de Sard.

L e m m e. *L'ensemble des valeurs critiques d'une fonction différentiable sur l'intervalle* [0, 1] *est de mesure nulle.*

◀ Partageons l'intervalle [0, 1] en N parties égales et repérons celles d'entre elles qui contiennent les points critiques. Si N est assez grand, la dérivée de la fonction n'est pas supérieure à C/N (C est une constante indépendante de N) sur chacune des parties repérées. Donc la longueur de l'image de chacune des parties repérées est $\leqslant C/N^2$. Recouvrons cette image par un intervalle de longueur $2C/N^2$. Nous obtenons un recouvrement de l'ensemble des valeurs critiques par des intervalles dont la longueur totale est $\leqslant 2C/N$. ▶

Considérons la famille de champs de vecteurs de paramètre ε sur le cercle, définie par $v(x, \varepsilon) = v(x) - \varepsilon$. Le point x est un point singulier dégénéré du champ associé à la valeur ε du paramètre si et seulement si ε est valeur critique de la fonction $v(x)$.

Or les valeurs critiques forment un ensemble de mesure nulle. Il existe donc des valeurs non critiques aussi petites que l'on veut. Fixons une valeur critique ε. Tous les points singuliers du champ correspondant à cette valeur du paramètre sont non dégénérés. ▶

F. Digression: théorème de Sard.

Soit $f: M \to N$ une application différentiable, M et N étant des variétés de dimension quelconque. On dira qu'un point de l'espace antécédent est *critique* si la dimension de l'image de la différentielle de l'application en ce point est inférieure à celle de l'espace de l'image. La valeur de l'application en un point critique est dite *valeur critique*.

T h é o r è m e. *L'ensemble des valeurs critiques de toute application assez lisse est de mesure nulle.*

◀1°. Si l'espace antécédent est de dimension nulle, le théorème est évident. S'il est de dimension un, le théorème est prouvé plus haut. Supposons que le théorème est valable pour tous les cas où l'espace antécédent est de dimension $m - 1$ et prouvons qu'il l'est pour un espace antécédent de dimension m.

2°. Partageons l'ensemble K des points critiques de l'application. Par définition, un *point méplat d'ordre r* est un point de l'espace antécédent en lequel

sont nulles toutes les dérivées partielles d'ordre 1, . . ., r. Notons K_r l'ensemble de tous les points méplats d'ordre r.

3°. Considérons tout d'abord l'ensemble des points critiques $K \setminus K_1$. Montrons que *l'ensemble correspondant de valeurs critiques* (c'est-à-dire l'ensemble $f(K \setminus K_1)$) *est de mesure nulle.*

En chaque point de $K \setminus K_1$ il existe une dérivée partielle, disons $\partial f_1/\partial x_1$, qui est non nulle. Au voisinage d'un tel point on peut prendre la fonction f_1 pour coordonnée locale dans l'espace antécédent à la place de x_1 et conserver $x_2, \ldots, x_m$. Dans ces coordonnées f se note comme une famille à un paramètre d'applications différentiables d'espaces de dimension $m - 1$ dans des espaces de dimension $n - 1$:

$$(f_1; x_2, \ldots, x_m) \longmapsto (f_1; f_2, \ldots, f_n).$$

Figeons la valeur c du paramètre f_1. L'application f induit une application f_c du plan $f_1 = c$ de dimension $m - 1$ de l'espace antécédent dans le plan $f_1 = c$ de dimension $n - 1$ dans l'espace image.

Les valeurs critiques de l'application f_c forment dans le plan $f_1 = c$ de l'espace image un ensemble de mesure de dimension $n - 1$, nulle par hypothèse de la récurrence (le théorème a été prouvé pour les antécédents de dimension $m - 1$). D'après le théorème de Fubini la mesure de dimension n de la réunion des ensembles de valeurs critiques des applications f_c par rapport à c est nulle.

Or l'image de l'ensemble des points critiques de l'ensemble $K \setminus K_1$, contenus dans le voisinage du point considéré, appartient à cette réunion. D'où il suit que la mesure de $f(K \setminus K_1)$ est nulle.

4°. Considérons l'ensemble $K_r \setminus K_{r+1}$ des points méplats d'ordre r. Montrons que *la mesure de l'ensemble correspondant de valeurs critiques de* $f(K_r \setminus K_{r+1})$ *est nulle.*

En tout point de l'ensemble $K_r \setminus K_{r+1}$ est non nulle l'une des dérivées partielles d'ordre $r + 1$, mettons, $\partial g/\partial x_1$, où g est l'une des dérivées partielles d'ordre r de f_1 dans des coordonnées locales convenables.

Au voisinage de l'un de ses points l'ensemble $K_r \setminus K_{r+1}$ est contenu dans une hypersurface $g = 0$, différentiable, de dimension $m - 1$. Les points de l'ensemble $K_r \setminus K_{r+1}$ sont critiques pour la restriction de f à cette hypersurface, puisque $df = 0$ sur K_r. Or, par hypothèse, la mesure de l'image de l'ensemble des valeurs critiques de la restriction de f à cette hypersurface est nulle. Donc la mesure de $f(K_r \setminus K_{r+1})$ est nulle.

5°. Considérons enfin l'ensemble K_r des points critiques d'ordre r du plan pour r assez grand. Prouvons que *la mesure de l'ensemble correspondant de valeurs critiques de* $f(K_r)$ *est nulle si r est assez grand.*

A cet effet, divisons un cube de dimension m de l'espace antécédent (après avoir fait choix de coordonnées locales) en N^m petits cubes égaux et repérons ceux d'entre eux qui contiennent des points de K_r. Le diamètre de l'image d'un cube repéré est alors $\leqslant c\,(1/N)^{r+1}$ (où la constante c est indépendante de N). Donc toutes les images des cubes repérés sont recouvertes par des cubes ouverts de mesure globale non supérieure à

$$c_1 N^m (1/N)^{n(r+1)},$$

même si tous les N^m cubes ont été repérés.

Ce nombre tend vers zéro pour $N \to \infty$, par suite mes $f(K_r) = 0$ pour $r > (m/n) - 1$.

Représentons l'ensemble K des points critiques comme la réunion des ensembles $K \setminus K_1$, $K_i \setminus K_{i+1}$ et K_r. Nous avons montré que l'image de chacun de ces ensembles était de mesure nulle. Donc l'ensemble des valeurs critiques est de mesure nulle. ▶

G. Systèmes structurellement stables sur une sphère de dimension deux.

L'étude des systèmes à espaces des phases de dimension supérieure à un passe d'abord par celle des points singuliers et des trajectoires fermées.

D é f i n i t i o n. Un point singulier d'un champ de vecteurs est *dégénéré* si le zéro est valeur propre du linéarisé du champ en ce point.

R e m a r q u e. Un point singulier non dégénéré ne disparaît pas par une petite déformation du champ, mais se déplace légèrement (d'après le théorème des fonctions implicites). Au contraire, un point singulier dégénéré bifurque généralement (se décompose en plusieurs points dégénérés) ou disparaît. Donc tous les points singuliers d'un système structurellement stable sont non dégénérés.

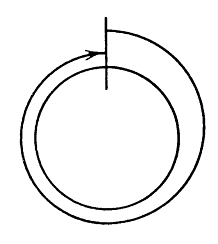


Fig. 72.

D é f i n i t i o n. Une trajectoire fermée (un cycle) d'un champ de vecteurs est *dégénérée* si 1 est valeur propre de la linéarisée de l'application de Poincaré. [On appelle application de Poincaré l'application d'une transversale au cycle dans elle-même, qui à tout point de la transversale, proche du cycle, associe le point d'intersection suivant de la trajectoire, issue de ce point avec cette transversale, cf. fig. 72].

R e m a r q u e. Un cycle non dégénéré ne disparaît pas par une petite déformation du champ, il se déplace légèrement (en vertu du

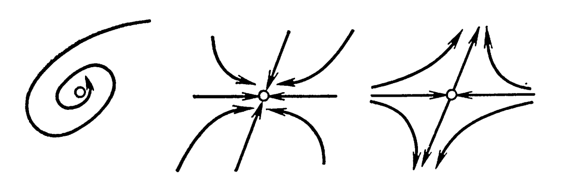


Fig. 73.

théorème des fonctions implicites). Au contraire, un cycle dégénéré bifurque (se décompose en plusieurs cycles non dégénérés) ou disparaît. Donc, tous les cycles d'un système structurellement stable sont non dégénérés.

Considérons un champ de vecteurs sur une surface plane. En dimension deux les points singuliers non dégénérés sont topologiquement soit des cols, soit des nœuds. Une trajectoire tendant vers un col pour $t \to \infty$ est dite séparatrice rentrante et pour $t \to -\infty$, séparatrice partante (fig. 73).

T h é o r è m e. *Un champ de vecteurs sur une sphère à deux dimensions définit un système structurellement stable si et seulement si sont réalisées les conditions suivantes:*

1) *Le champ possède un nombre fini de points singuliers.*
2) *Tous les points singuliers sont non dégénérés.*
3) *Aucune séparatrice partante du col n'est rentrante.*
4) *Le champ possède un nombre fini de trajectoires fermées.*
5) *Toutes les trajectoires fermées sont des cycles non dégénérés.*

R e m a r q u e. La démonstration de l'instabilité structurelle d'un système dans le cas où l'une au moins des cinq conditions énumérées est violée n'est pas compliquée (cf. fig. 74). La démonstration du fait que la stabilité structurelle découle des conditions 1) à 5) est plus compliquée. Elle est effectuée en détail

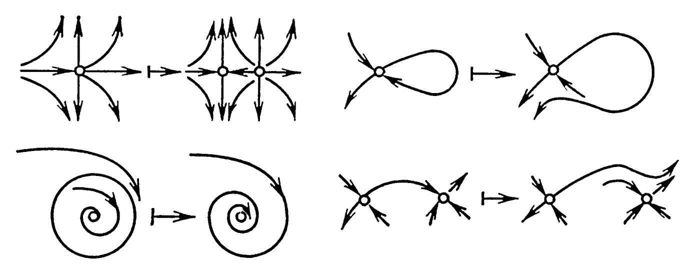


Fig. 74.

par De Baggis (H. F. D e B a g g i s, Dynamical systems with stable structures, Contr. Theory Nonlinear Oscillation, 2 (1952), 37-59; M. M. P e i x o t o, Structural stability on two-dimensional manifolds, Topology, 1 (1962), 101-120; 2 (1963), 179-180).

Pour les systèmes structurellement stables sur le plan voir également H. F. D e B a g g i s Dynamical systems with stable to the theory of nonlinear oscillations, Prinston University Press, t. II, 1952, 37-59; M. C. P e i x o t o, M. M. P e i x o t o, Structural stability in the plane with enlarged boundary conditions, Anais da Acad. Brasileira de Ciências, **31**, 2 (1959), 135-160.

T h é o r è m e. *Les champs de vecteurs structurellement stables forment un ensemble partout dense dans l'espace des champs de vecteurs sur une sphère à deux dimensions.*

◀ Ce théorème est une conséquence du précédent. ▶

R e m a r q u e. Cette proposition est valable pour les champs de vecteurs sur un disque non tangents à la frontière de ce disque.

§ 11. Equations différentielles sur le tore

On expose ici la théorie des champs de vecteurs sans points singuliers sur un tore à deux dimensions, qui est l'œuvre de H. Poincaré et A. Denjoy, et on décrit tous les champs structurellement stables.

A. Tore à deux dimensions.

On appelle *tore à n dimensions* le produit direct de n cercles. Le tore à deux dimensions $T^2 = S^1 \times S^1$ peut être représenté comme un carré

$$\{x,\ y\colon 0 \leqslant x \leqslant 2\pi,\quad 0 \leqslant y \leqslant 2\pi\}$$

dont on a collé les côtés opposés (on identifie les points $(0, y)$ et $(2\pi, y)$ ainsi que $(x, 0)$ et $(x, 2\pi)$ fig. 75).

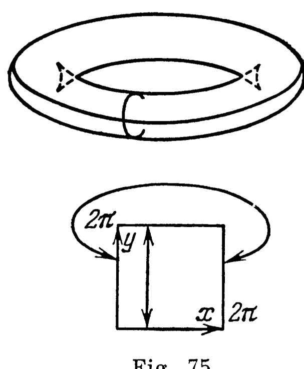


Fig. 75

Le tore peut être également défini comme le quotient de l'ensemble des classes d'équivalence du groupe $\mathbb{R}^2$ par le sous-groupe $2\pi\mathbb{Z}^2$ des vecteurs entiers multipliés par 2π :

$$T^2 = \mathbb{R}^2/2\pi\mathbb{Z}^2 = \{(x,\ y) \in \mathbb{R}^2 \bmod 2\pi\}.$$

Le plan $\mathbb{R}^2$ recouvre donc le tore par un difféomorphisme local. Le revêtement $\mathbb{R}^2 \to T^2$ (fig. 76) permet de transporter toute figure

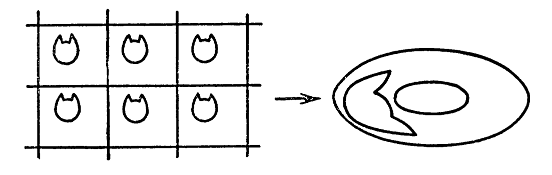


Fig. 76

du tore sur le plan (où elle est indéfiniment reproduite). Les fonctions différentiables sur le tore sont associées à des fonctions différentiables 2π-périodiques sur le plan.

A toute trajectoire fermée du plan est associée une trajectoire fermée du tore. La réciproque est fausse : les trajectoires fermées du tore sont associées aussi bien à des trajectoires fermées du plan qu'aux applications $\varphi\colon [0, 1] \to \mathbb{R}^2$, telles que $\varphi(0) = \varphi(1) \bmod 2\pi$.

Si les coordonnées $\varphi(1) - \varphi(0)$ sont égales à $(2\pi p, 2\pi q)$ on dit que la trajectoire *se boucle après avoir effectué p tours sur un parallèle et q tours sur un méridien.*

B. Champs de vecteurs sur le tore.

Tout champ de vecteurs sur le tore définit sur le plan un champ 2π-périodique par rapport aux deux coordonnées. Inversement, à tout champ 2π-périodique par rapport aux deux coordonnées sur le plan est associé un champ de vecteurs sur un tore.

Exemple. L'équation

$$\dot{x} = \alpha, \qquad \dot{y} = \beta,$$

où α et β sont des constantes, définit un champ de vecteurs sans points singuliers sur le tore.

Théorème. *Les trajectoires de l'équation sur le tore sont fermées si le quotient $\lambda = \beta/\alpha$ est rationnel, et partout denses s'il est irrationnel.*

◀ 1) Soit $\lambda = p/q$. La trajectoire passant par le point (x_0, y_0) a pour équation $y - y_0 = (x - x_0) p/q$. Si $x - x_0 = 2\pi q$, alors $y - y_0 = 2\pi p$ et, par suite, $(x, y) = (x_0, y_0) \bmod 2\pi$ est une trajectoire fermée.

2) Nous allons montrer que la trajectoire (dans le cas où λ est irrationnel) est uniformément distribuée sur le tore, c'est-à-dire passe dans chaque partie du tore *) une durée de temps proportionnelle à l'aire de cette partie. D'où il résultera, en particulier, que la trajectoire est partout dense dans le tore. ▶

C. Répartition uniforme.

Voici la définition générale de la répartition uniforme.

Soit v un champ de vecteurs sur une variété différentiable compacte M dont l'élément de volume est fixé (par exemple un champ sur un tore d'élément de surface $dx\, dy$). On notera $\mu(D)$ le volume (resp. la surface) d'un domaine D.

Considérons la solution φ de l'équation $\dot{x} = v(x)$ vérifiant une condition initiale z. Désignons par $\tau(D, T, z)$ la mesure de l'ensemble des valeurs de $t \in [0, T]$ pour lesquelles $\varphi(t)$ appartient à D.

Définition. Les solutions de l'équation $\dot{x} = v(x)$ sont *uniformément réparties* si pour tout domaine D de bord différentiable par morceaux

$$\lim_{T\to\infty} \frac{\tau(D, T, z)}{T} = \frac{\mu(D)}{\mu(M)}$$

*) Par partie du tore on entend un domaine mesurable-Jordan, par exemple un domaine à bord différentiable par morceaux.

Théorème. *Les solutions de l'équation* $\dot{x} = \alpha$, $\dot{y} = \beta$ *sont uniformément réparties sur le tore si* β/α *est irrationnel.*

La répartition uniforme peut être également définie en termes de *moyennes temporelles* de fonctions.

Soit f une fonction sur M (d'une façon générale à valeurs complexes).

Définition. La limite

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(g^t z)\, dt = \bar{f}(z)$$

s'appelle moyenne temporelle de la fonction f (g^t est le flot).

Remarque. Il est évident que cette limite n'existe pas toujours, et si elle existe, elle dépend généralement du point initial.

Le théorème de la répartition uniforme sur le tore est une conséquence du

Théorème (de la coïncidence des moyennes). *Si* $\lambda = \beta/\alpha$ *est irrationnel, la moyenne temporelle de toute fonction continue (ou à la rigueur intégrable-Riemann)* $f: T^2 \to \mathbb{C}$ *le long des solutions de l'équation* $\dot{x} = \alpha$, $\dot{y} = \beta$ *existe, est indépendante du point initial et coïncide avec la moyenne spatiale :*

$$f_0 = \frac{1}{4\pi^2} \oiint f\, dx\, dy.$$

◀ Pour déduire le théorème de la répartition uniforme à partir de celui-ci il suffit de prendre pour f la fonction caractéristique de l'ensemble D (qui est égale à l'unité sur D et nulle en dehors de D). ▶

D. Démonstration du théorème de coïncidence des moyennes. Soit ω le vecteur de coordonnées (α, β). La solution vérifiant la condition initiale z s'écrit alors $z + \omega t$. Le théorème affirme que

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(z + \omega t)\, dt = \frac{1}{4\pi^2} \oiint f(z)\, dx\, dy.$$

◀ On observera que parmi les fonctions sur le tore il existe des harmoniques, c'est-à-dire des fonctions de la forme $e^{i(k,z)}$, où k est un vecteur entier. Pour les harmoniques on vérifie le théorème par un calcul immédiat de l'intégrale. Supposons $k \neq 0$. On a alors

$$\int_0^T e^{i(k,\, z+\omega t)}\, dt = e^{ikz} \int_0^T e^{i(k,\, \omega)t}\, dt = \frac{e^{i(k,\, z)}}{i(k,\, \omega)} [e^{i(k,\, \omega)T} - 1].$$

La fonction entre crochets est bornée, par suite la moyenne temporelle de l'harmonique d'indice $k \neq 0$ est nulle. La moyenne spa-

tiale l'est également. Pour $k = 0$ l'harmonique est égale à l'unité. Les deux moyennes de l'unité sont égales à l'unité. Ceci prouve la coïncidence des moyennes des harmoniques.

Cette coïncidence entraîne celle des moyennes des polynômes trigonométriques : la moyenne d'une combinaison linéaire est égale à la combinaison linéaire des moyennes ayant les mêmes coefficients. En particulier, le théorème est prouvé pour $f = \cos(k, z)$ et $\sin(k, z)$.

Prouvons maintenant le théorème pour les fonctions réelles : ce théorème sera valable aussi pour les fonctions complexes (en raison de la linéarité des moyennes). Encadrons la fonction f par des fonctions continues P et Q telles que

$$P < f < Q, \quad \oiint (Q - P)\, dx\, dy/4\pi^2 < \varepsilon$$

(un tel encadrement pour tout $\varepsilon > 0$ est caractéristique des fonctions intégrables-Riemann). Approchons ensuite les fonctions P et Q par des polynômes trigonométriques p et q tels que $|p - P| < \varepsilon$, $|q - Q| < \varepsilon$.

Désignons par p_0 et q_0 les termes libres de ces polynômes. Les nombres p_0 et q_0 sont les moyennes temporelles et spatiales des polynômes p et q (puisque les moyennes temporelles et spatiales des polynômes trigonométriques sont confondues). Donc, la moyenne spatiale f_0 de la fonction f est comprise entre p_0 et q_0 :

$$p_0 < f_0 < q_0, \quad q_0 - p_0 < \varepsilon.$$

Désignons par p_T, f_T et q_T les moyennes de p, f et q dans le temps T :

$$p_T(z) = \frac{1}{T} \int_0^T p(z + \omega t)\, dt, \ \text{etc.}$$

Alors $p_T(z) < f_T(z) < q_T(z)$, $\forall\, T$, et pour T assez grand on a

$$|p_T(z) - p_0| < \varepsilon, \quad |q_T(z) - q_0| < \varepsilon.$$

Donc pour T assez grand

$$|f_T(z) - f_0| < 2\varepsilon. \ \blacktriangleright$$

E. Quelques corollaires.

1°. La dimension deux du tore n'a à aucun moment intervenu dans ce qui précède. Soit l'équation $\dot{z} = \omega$, $z \in T^n$ sur un tore à n dimensions. Le vecteur fréquence ω est dit *résonnant* si existe un vecteur k entier non nul tel que $(\omega, k) = 0$.

Si le vecteur ω n'est pas résonnant, les moyennes temporelles et spatiales de fonctions continues (ou à la rigueur intégrables-Riemann) sont confondues et les solutions sont uniformément réparties.

2°. Le théorème de la répartition uniforme nous dit que le premier chiffre du nombre 2^n est plus souvent un 7 qu'un 8. Plus exactement, désignons par $N_k(n)$ le nombre des nombres naturels $m \leqslant n$ pour lesquels 2^m commence par le chiffre k. Alors

$$\lim_{n\to\infty} \frac{N_7(n)}{N_8(n)} = \frac{\lg 8 - \lg 7}{\lg 9 - \lg 8}.$$

3°. Le théorème de répartition uniforme doit son existence au problème suivant de Lagrange: trouver $\omega = \lim_{t\to\infty} \frac{1}{t} \arg f(t)$, où $f(t) = \sum_{k=1}^{n} a_k e^{i\omega_k t}$.

Exhibons la réponse pour le vecteur non résonnant $\omega = (\omega_1, \ldots, \omega_n)$. Soit $n = 3$. Si l'on peut construire un triangle de côté (a_1, a_2, a_3), alors $\omega = \Sigma\, \alpha_k \omega_k / \pi$, où α_k est l'angle opposé au côté a_k.

Pour n quelconque, la réponse est donnée par la moyenne pondérée des fréquences $\omega_k : \omega = \Sigma\, W_k\, \omega_k$. Les poids W_k se calculent comme suit. Soit $W(a_1, \ldots, a_s; b)$ la probabilité que la distance entre l'origine et l'extrémité d'une ligne polygonale plane de s segments de longueur $a_1, \ldots, a_s$ aléatoirement orientés soit inférieure à b. Alors $W_k = W(\hat{a}_k; a_k)$ ($\hat{a}_k$ représente tous les nombres a_i à l'exception de a_k).

Une démonstration figure dans l'article de H. Weyl Mean Motion, Amer. J. Math. 60 (1938), 889-896.

Voici comment Lagrange a été conduit à étudier le problème précédent (dit problème du mouvement moyen). Considérons le vecteur joignant le Soleil au centre de l'ellipse sur laquelle se déplace une planète (on l'appelle vecteur de Laplace). En première approximation de la théorie des perturbations, l'évolution du vecteur de Laplace sous l'influence de la gravitation universelle des planètes a la forme d'un mouvement d'une somme de vecteurs en rotation uniforme (leur nombre est égal à celui des planètes).

Si l'on calcule les fréquences ω_k et les amplitudes a_k des planètes du système solaire, on constate que pour toutes les planètes, sauf la Terre et Vénus, l'une des amplitudes a_k est plus grande que la somme de toutes les autres. C'est pourquoi Lagrange a réussi à trouver le mouvement moyen des périhélies de toutes les planètes à l'exception de la Terre et de Vénus. Dans le cas de ces deux dernières quelques termes possèdent environ la même amplitude. Ce problème n'a été résolu qu'au vingtième siècle dans les travaux de Bol, Serpinski et H. Weyl.

F. Application de Poincaré et fonction angulaire.

Soit l'équation différentielle générale sur le tore

$$\dot{z} = \omega(z), \qquad z \in T^2.$$

Supposons que le champ ω ne possède pas de points singuliers et de plus $\omega_1 \neq 0$. [Si le champ ne possède pas de points singuliers et de cycles, sa première composante est partout différente de zéro dans

un système de coordonnées convenable (voir C. L. Siegel, Note on differential equations on the torus, Ann. Math. 46, 3 (1945). 423-428); il est aisé de construire un champ sans points singuliers mais avec des cycles n'autorisant pas un tel système de coordonnées).]

Nous sommes conduits à étudier maintenant les courbes intégrales d'une équation non autonome à second membre bipériodique

$$\frac{dy}{dx} = \lambda(x, y), \qquad \lambda = \omega_2/\omega_1.$$

Toutes les solutions de cette équation sont indéfiniment prolongeables puisque le second membre est borné.

D é f i n i t i o n. On appelle *application de Poincaré* pour l'équation étudiée sur le tore une application A de l'axe y dans lui-même associant à tout point initial $(0, y_0)$ la valeur de la solution vérifiant cette condition initiale pour $x = 2\pi$ (fig. 77).

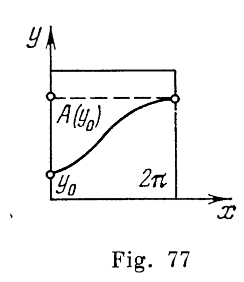


Fig. 77

L'application de Poincaré est dérivable (d'après le théorème de dérivabilité de la solution par rapport aux conditions initiales) et périodique $A(y + 2\pi) = A(y) + 2\pi$; l'application inverse A^{-1} est aussi dérivable. Donc, A définit un difféomorphisme du cercle sur lui-même. On peut se représenter l'application de Poincaré comme un difféomorphisme du méridien du tore dans lui-même associant à tout point du méridien le point d'intersection suivant de la courbe intégrale passant par ce point avec ce méridien.

L'étude des propriétés des courbes intégrales sur le tore se ramène ainsi à celles des difféomorphismes sur le cercle. Supposons par exemple qu'un difféomorphisme du cercle possède un point fixe. Il lui correspond alors sur le tore une courbe intégrale fermée. La réciproque n'est pas vraie: exemple, la rotation du cercle d'un angle π. Pour qu'une courbe intégrale passant par un point donné du méridien du tore soit fermée il faut et il suffit que ce point soit point périodique du difféomorphisme, c'est-à-dire qu'il se transforme en lui-même par une puissance du difféomorphisme.

L'application de Poincaré définit un difféomorphisme du cercle conservant l'orientation. On peut donc l'écrire

$$Ay = y + a(y), \text{ où } a(y + 2\pi) = a(y),\ a'(y) > -1.$$

La fonction a est dite *fonction angulaire.*

G. Nombre de rotation.

Le nombre de rotation caractérise la pente moyenne des courbes intégrales de l'équation sur le tore; pour l'équation élémentaire à coefficients constants $dy/dx = \lambda$ ce nombre est λ.

D é f i n i t i o n. On appelle *nombre de rotation* de l'équation $dy/dx = \lambda(x, y)$ sur le tore le nombre

$$\mu = \lim_{x\to\infty} \frac{\varphi(x)}{x},$$

où φ est une solution sur le plan.

Entre le nombre de rotation et la fonction angulaire on a la relation suivante:

$$\mu = \frac{1}{2\pi} \lim_{k\to\infty} \frac{a(y)+a(Ay)+\ldots+a(A^{k-1}y)}{k}.$$

Dans cette forme la définition se transpose à tout difféomorphisme du cercle préservant l'orientation.

T h é o r è m e. *La limite figurant dans la définition du nombre de rotation existe et ne dépend pas du point initial; elle est rationnelle si et seulement si une puissance du difféomorphisme possède un point fixe (c'est-à-dire lorsque l'équation différentielle possède une trajectoire fermée).*

◀ 1°. Considérons l'angle de rotation du point y par la puissance k d'un difféomorphisme. Désignons-le par

$$a_k(y) = a(y) + a(Ay) + a(A^2y) + \ldots + a(A^{k-1}y).$$

On a la relation

$$|a_k(y_1) - a_k(y_2)| < 2\pi,$$

où y_1 et y_2 sont des points quelconques.

En effet, l'inégalité est réalisé pour $|y_1 - y_2| < 2\pi$, puisque les images de segments de longueur 2π par les applications A et A^k de la droite sont des segments de longueur 2π. Or la fonction a_k est 2π-périodique, donc on peut modifier y_2 d'un multiple entier de 2π, de sorte que $a_k(y_2)$ ne varie pas et la distance de y_1 à y_2 devienne inférieure à 2π.

2°. Soit m_k un entier tel que

$$2\pi m_k \leqslant a_k(0) < 2\pi(m_k + 1).$$

Montrons que *pour tout y et tout l entier*

$$\left|\frac{a_{kl}(y)}{2\pi kl} - \frac{m_k}{k}\right| < \frac{2}{k}.$$

En effet, en vertu du 1°, on a

$$|a_k(y) - 2\pi m_k| < 4\pi, \quad \forall y,$$

donc

$$\left|\frac{a_k(y)}{2\pi kl} - \frac{m_k}{k}\right| < \frac{2}{k}.$$

Or $a_{kl}(y)/2\pi kl$ est la moyenne arithmétique des quantités $a_k(y_i)/2\pi k$, où $y_i = A^i y$, $i = 0, \ldots, l-1$.

3°. Soit σ_k l'intervalle $\left[\frac{m_k-2}{k}, \frac{m_k+2}{k}\right]$. Nous avons montré que $a_k(y)/2\pi kl$ *appartient* à σ_k *pour tous les* l. Montrons que tous *les intervalles* σ_k *se coupent.*

En effet, $a_{kl}(y)/2\pi kl$ appartient aussi bien à σ_k qu'à σ_l.

4°. Ainsi *les intervalles* σ_k *tendent vers zéro et s'intersectent deux à deux.* Ils possèdent donc un point d'intersection unique qui n'est autre que le nombre de rotation. Nous avons montré que la limite définissant le nombre de rotation existe et ne dépend pas du point initial.

5°. Supposons que A^q possède un point fixe y sur le cercle ; l'image de ce point sur la droite par l'application A^q se déplace d'un multiple entier de 2π, c'est-à-dire $a_q(y) = 2\pi p$. Dans ce cas, $a_{ql}(y) = 2\pi pl$, $\forall l$, donc *le nombre de rotation* $\mu = p/q$ *est rationnel.*

6°. Soit $\mu = p/q$. Si pour tous les y on a $a_q(y) > 2\pi p$, alors pour un $\varepsilon > 0$ on aura $a_q(y) > 2\pi p + \varepsilon$ pour tous les y.

Or à ce moment $\mu > p/q$. Si $a_q(y) < 2\pi p$, $\forall y$, on aurait $\mu < p/q$. Donc $a_q - 2\pi p$ change de signe, par suite existe y tel que $a_q(y) = 2\pi p$. ▶

R e m a r q u e. Si le nombre de rotation μ est irrationnel, les points $(y, Ay, A^2y, \ldots, A^N y)$ sont disposés sur le cercle pour tout y dans le même ordre que dans le cas d'une rotation de $2\pi\mu$. En effet, $a_q(y) > 2\pi p$ si et seulement si $\mu > p/q$.

A noter encore que le nombre de rotation d'une équation sur le tore *dépend* du choix du cercle transversal à la trajectoire (l'axe y dans nos notations).

H. Equations structurellement stables sur le tore.

L'équation élémentaire sur le tore $\dot{z} = \omega$ est structurellement instable pour les valeurs aussi bien résonnantes que non résonnantes de ω.

T h é o r è m e 1. *L'équation différentielle sur le tore* $dy/dx = \lambda(x, y)$ *est structurellement stable si et seulement si le nombre de rotation est rationnel et toutes les solutions périodiques non dégénérées* *).

◀ Ce théorème est une conséquence d'une proposition démontrée plus bas sur les difféomorphismes du cercle conservant l'orientation. ▶

D é f i n i t i o n. On appelle *cycle d'ordre* q d'un difféomorphisme $A : M \to M$ l'ensemble de q points $(y, Ay, \ldots, A^{q-1}y)$ sous réserve qu'ils soient tous distincts et $A^q y = y$. Le cycle est non dégénéré si son point y est un point fixe non dégénéré de l'application A^q (c'est-

*) A. Mayer, Transformations grossières d'un cercle dans un cercle, Outcheny zapiski, GGU, 12 (1939), 215-229 ; V. Pliss, Sur la grossièreté des équations différentielles définies sur le tore, Vestnik LGU, ser. mat. 13, 3 (1960), 15-23.

à-dire 1 n'est pas valeur propre de la dérivée de l'application A^q au point y).

R e m a r q u e. Les dérivées de l'application A^q en divers points d'un cycle sont semblables, donc les points d'un cycle sont ou tous dégénérés ou tous non dégénérés.

T h é o r è m e 2. *Un difféomorphisme du cercle préservant l'orientation est structurellement stable si et seulement si le nombre de rotation est rationnel et tous ses cycles sont non dégénérés. Les difféomorphismes structurellement stables forment un ensemble partout dense dans l'espace C^2 des difféomorphismes du cercle préservant l'orientation.*

Donc, les difféomorphismes génériques à nombre de rotation rationnel sont construits de façon assez simple. Le type topologique de l'application est déterminé par le nombre de cycles qui doit être pair (par suite de l'alternance des points des cycles stables et des cycles instables). L'ordre de tous les cycles est q si le nombre de rotation $\mu = p/q$. L'ordre des points d'un cycle sur le cercle est le même que pour une rotation d'angle $2\pi\mu$.

Le théorème 2 est prouvé plus bas au point J. La démonstration n'est pas compliquée si l'on admet le théorème non trivial suivant, dû à Denjoy (1932).

T h é o r è m e 3. *Si un difféomorphisme du cercle de classe C^2 possède un nombre de rotation μ irrationnel, il est topologiquement équivalent à une rotation du cercle d'un angle de $2\pi\mu$.*

La théorie précédente est due à Poincaré (1885); Poincaré a énoncé le théorème de Denjoy sous forme d'hypothèse (pour les équations dont le second membre est un polynôme trigonométrique). Denjoy a exhibé également des exemples prouvant que C^2 ne pouvait pas être remplacé par C^1.

I. Démonstration du théorèmes de Denjoy.

◀ 1°. Les points . . ., $A^{-1}y$, y, Ay, A^2y, . . . de l'orbite de l'application A sont disposés sur le cercle dans le même ordre que les points de l'orbite d'une rotation d'un angle de $2\pi\mu$ (voir point G). Donc *pour prouver le théorème il suffit d'établir que l'orbite de l'application A est partout dense dans le cercle.* En effet, on obtiendra un homéomorphisme du cercle associant une rotation à A en prolongeant en continuité une application qui associe aux points . . ., $A^{-1}y$, y, Ay, . . . les points correspondants de l'orbite de la rotation.

2°. *Si un arc de cercle ne contient pas de points de l'orbite de A, alors les images de cet arc par les puissances de A ne s'intersectent pas deux à deux.* Considérons en effet le plus grand arc comprenant l'arc donné mais pas de points de l'orbite. Toutes ses images sont des arcs maximaux. Les extrémités de l'arc maximal appartiennent à l'adhérence de l'orbite. Donc les extrémités des arcs maximaux ne peuvent être contenues dans les arcs maximaux. Par suite, deux arcs maximaux ayant un point commun sont nécessairement con-

fondus. Or si un arc maximal est confondu avec son image, cela veut dire que son extrémité appartient au cycle, ce qui contredit l'irrationnalité de μ.

3°. La somme des longueurs des images d'un arc maximal est bornée. Par suite, les longueurs des images successives d'un tel arc aussi bien par A^N que A^{-N} tendent vers zéro pour $N \to \infty$. Donc *les intégrales des jacobiens des itérations positives et négatives de A sur l'arc maximal tendent vers zéro : si l'on note*

$$u_N = \prod_{i=0}^{N-1} \frac{dA}{dy}(A^i y), \quad v_N = \prod_{i=0}^{N-1} \frac{dA^{-1}}{dy}(A^{-i} y),$$

alors pour $N \to \infty$ on a

$$\int u_N dy \to 0, \quad \int v_N dy \to 0$$

(l'intégration est étendue à l'arc maximal).

4°. Soit $(\alpha_0, \alpha_1, \ldots)$ une suite de points de l'orbite d'une rotation d'angle $2\pi\mu$. *Supposons que α_q est celui des points $(\alpha_1, \ldots, \alpha_q)$ qui est le plus proche de α_0. Alors les points $\alpha_q, \ldots, \alpha_{2q-1}$ alternent avec les points $\alpha_0, \ldots, \alpha_{q-1}$.*

En effet, soit l'arc (α_s, α_{q+s}), $s < q$, de longueur δ égale à la distance de α_0 à α_q. Supposons que α_r est situé sur cet arc. Si $r < s$, l'arc $(\alpha_{s-r}, \alpha_{s-r+q})$ contient α_0, donc la distance de α_{s-r} à α_0 est inférieure à δ contrairement au choix de α_q. Si $r > s$, alors l'arc (α_0, α_q) contient α_{r-s}, donc $r - s > q$. Or la distance de α_0 à α_{r-s-q} est alors inférieure à δ. Donc, l'arc (α_s, α_{q+s}) ne contient pas de points α_r, $r < 2q$, c.q.f.d.

5°. Soient les points $(y, Ay, \ldots, A^{q-1}y)$ et $(A^{-1}y, \ldots, A^{-q}y)$. Ces deux ensembles de points alternent (cf. 4°). Donc quels que soient f, fonction à variation bornée sur le cercle, un point y et q, défini au 4°, *la quantité*

$$\sum f(A^i y) - \sum f(A^{-j} y), \quad 0 \leqslant i < q, \ 0 < j \leqslant q$$

est encadrée par des constantes indépendantes de y et q.

6°. Prenons pour f la fonction $\ln (dA/dy)$. Cette fonction est à variation bornée puisque A appartient à la classe C^2. Donc *la quantité*

$$\prod_{i=0}^{q-1} \frac{dA}{dy}(A^i y) \Big/ \prod_{j=1}^{q} \frac{dA}{dy}(A^{-j} y) = u_q v_q$$

est encadrée par des constantes positives indépendantes de y et q (sous réserve que q soit choisi comme au 4°).

7°. La contradiction avec le 3° achève la démonstration du théorème: l'inégalité de Schwarz appliquée à $\sqrt{u_q}$ et $\sqrt{v_q}$ donne

$$\left(\int \sqrt{u_q v_q}\, dy\right)^2 \leqslant \int u_q\, dy \int v_q\, dy. \quad \blacktriangleright$$

J. Démonstration du théorème des difféomorphismes structurellement stables du cercle.

◀ 1°. *Soient donnés deux difféomorphismes du cercle conservant l'orientation ayant le même nombre rationnel de rotation et le même nombre de cycles. Si tous les cycles sont non dégénérés, il existe un homéomorphisme associant un difféomorphisme à l'autre.*

Pour prouver cette proposition il faut d'abord associer aux points d'un cycle stable d'un difféomorphisme un point d'un cycle stable quelconque du second, ensuite du cycle instable voisin et ainsi de suite pour tous les cycles (on rappelle que l'ordre des points d'un cycle sur le cercle est le même que pour une rotation). On peut poursuivre cette correspondance aux intervalles adjacents en se servant du lemme suivant dont la démonstration ne présente aucune difficulté.

Deux homéomorphismes quelconques d'un intervalle sur lui-même, ne possédant pas de point fixe sont topologiquement conjugués.

2°. Si le nombre de rotation est rationnel et tous les cycles non dégénérés, une petite déformation préserve le nombre de rotation, le nombre et la non-dégénérescence des cycles (d'après le théorème des fonctions implicites). Donc tout *difféomorphisme à nombre de rotation rationnel et à cycle non dégénéré est structurellement stable* (voir 1°).

3°. Si un difféomorphisme possède un cycle dégénéré, par une petite déformation de ce difféomorphisme au voisinage des points de ce cycle, on peut modifier le nombre de cycles. Donc tout *difféomorphisme à cycle dégénéré est structurellement instable.*

4°. *Si le nombre de rotation est irrationnel on peut le modifier par une variation aussi petite que l'on veut du difféomorphisme.* Soit, en effet, $y \mapsto A(y) + \varepsilon$, $\varepsilon > 0$, le difféomorphisme perturbé. Le théorème de Denjoy nous dit que dans un système de coordonnées non différentiable $z \mapsto z + 2\pi\mu + \varphi(z)$, $\varphi > 0$. Donc le nombre de rotation du difféomorphisme perturbé est $> \mu$. Ainsi, *tout difféomorphisme à nombre de rotation irrationnel est structurellement instable.*

5°. *Le nombre de rotation est une fonction continue de difféomorphisme* En effet, $\mu < p/q$ si et seulement si tous les points se déplacent d'une distance inférieure à $2\pi p$ par la puissance q du difféomorphisme. Cette propriété reste valable pour une petite variation du difféomorphisme.

6°. *Les difféomorphismes à nombres rationnels de rotation forment un ensemble partout dense.* Ceci découle du 4°, 5° et de la densité de l'ensemble des nombres rationnels.

7°. *Tous les cycles d'un difféomorphisme à nombre de rotation rationnel peuvent être rendus non dégénérés par une variation aussi petite que l'on veut des difféomorphismes.*

En effet, une déformation aussi petite que l'on veut au voisinage d'un cycle rend ce dernier non dégénéré. Soit γ l'un des arcs découpés sur le cercle par un cycle non dégénéré. Définissons une fonction différentiable φ égale à 1 sur γ en dehors d'un petit voisinage des extrémités de γ, et nulle en dehors de γ. Posons $A_\varepsilon(y) = A(y) + \varepsilon\varphi(y)$. Le nombre de rotation de ce difféomorphisme n'a pas changé, puisque le cycle s'est conservé. Soit q l'ordre du cycle. Alors $A_\varepsilon^q(y)$ est confondu avec $A^q(y) + \varepsilon$ sur l'arc $A\gamma$ en dehors des extrémités de l'arc $A\gamma$.

En appliquant le lemme de Sard à la fonction $A^q(y) - y$ sur $A\gamma$ on s'assure que tous les points fixes de A_ε^q sur γ sont non dégénérés pour presque tous les ε. Or chaque cycle de l'application A possède un représentant sur l'arc $A\gamma$. Donc, tous les cycles de l'application A_ε sont non dégénérés. ▶

K. Discussion.

1°. Les théorèmes précédents donnent l'impression que le difféomorphisme « général » du cercle possède un nombre de rotation rationnel et que les difféomorphismes à nombre irrationnel de rotation constitue une exception. Les expériences numériques conduisent toutefois d'une façon générale à des orbites partout denses (du moins par l'allure). Pour expliquer ce phénomène considérons la famille des difféomorphismes

$$A_{\alpha,\varepsilon}: y \mapsto y + \alpha + \varepsilon \sin y, \quad \alpha \in [0, 2\pi], \quad \varepsilon \in [0, 1].$$

On représentera chaque difféomorphisme par un point du plan (α, ε). L'ensemble des difféomorphismes à nombre de rotation $\mu = p/q$ est limité (un calcul simple le prouve) par un couple de courbes différentiables et s'approche de l'axe $\varepsilon = 0$ en une cuspide d'autant plus étroite que q est grand. La réunion de ces ensembles est partout dense. Cependant la mesure de l'ensemble des points du plan des paramètres pour lesquels le nombre de rotation est rationnel dans le domaine $0 \leqslant \varepsilon \leqslant \varepsilon_0$, $0 \leqslant \alpha \leqslant 2\pi$ est petite en regard de la mesure de ce domaine (fig. 78).

Donc, un difféomorphisme pris au hasard dans la famille avec ε petit possède un nombre de rotation irrationnel avec une probabilité voisine de 1.

Bien plus, ceci est valable pour toute famille analytique ou suffisamment de fois différentiable de difféomorphismes proches des rotations, par exemple pour la famille $y \mapsto y + \alpha + \varepsilon\, a(y)$, a étant une fonction analytique quelconque: pour ε petit les orbites sont partout denses sur le cercle avec une probabilité voisine de 1 et le nombre de rotation est irrationnel.

Donc le point de vue de la stabilité structurelle n'est pas toujours l'unique approche de la notion de système générique. L'approche métrique indiquée plus haut convient souvent mieux à la description du comportement du système observé.

2°. D'après le théorème de Denjoy, une application différentiable à nombre de rotation irrationnel est topologiquement équivalente à une rotation. Il se pose la question de savoir si une application est différentiablement équivalente à une rotation.

La réponse est négative dans le cas où le nombre de rotation se laisse approcher anormalement vite par des nombres rationnels

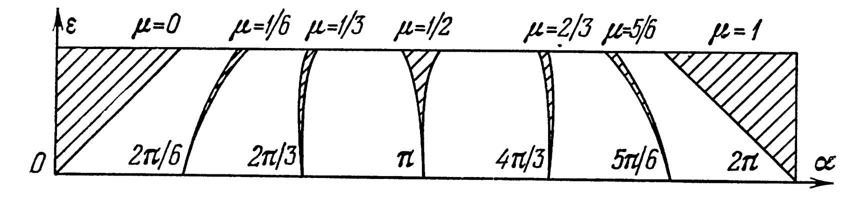


Fig. 78

(A. Finzi). Le problème de l'équivalence différentiable à une rotation se ramène à celui de la continuité d'une mesure invariante de l'application. Si le nombre de rotation est rationnel, la mesure est concentrée en des points isolés. Si le nombre de rotation se laisse approcher très vite par des nombres rationnels ayant des dénominateurs pas trop élevés, la mesure invariante se laisse tellement vite approcher par des mesures concentrées en des points isolées qu'elle ne peut même plus être absolument continue par rapport à la mesure de Lebesgue. Donc, les homéomorphismes ne peuvent être remplacés par des difféomorphismes dans le théorème de Denjoy.

3°. Dans l'approche métrique un nombre μ pris au hasard est presque sûrement irrationnel et, en outre, ne se laisse pas approcher assez rapidement par des nombres rationnels ayant de petits dénominateurs. Par exemple, pour tout $\varepsilon > 0$ existe presque sûrement $C > 0$ tel que

$$\left|\mu - \frac{p}{q}\right| < \frac{C}{q^{2+\varepsilon}}$$

quels que soient les entiers positifs p et q. D'où l'hypothèse que le fait discuté dans 2° n'a lieu qu'avec la probabilité zéro. Enonçons deux résultats allant dans ce sens.

Théorème. *Pour presque tout nombre de rotation μ chaque difféomorphisme (au moins de la classe C^3) du cercle de nombre de rotation μ est différentiablement équivalent à une rotation d'angle $2\pi\mu$* (M. Herman, 1976).

Le terme « presque partout » signifie que la mesure de Lebesgue de l'ensemble des nombres de rotation exclus est nulle.

Le théorème d'Herman fait suite à un théorème analogue pour les applications proches d'une rotation et au résultat suivant (démontré dans le cas analytique en 1959, et dans le cas différentiable par J. Moser en 1962).

Théorème. *Dans une famille suffisamment lisse de difféomorphismes* $A_{\alpha,\varepsilon}$: $y \mapsto y + \alpha + \varepsilon a(y)$, *la mesure de l'ensemble des couples* (α, ε) *dans le domaine* $0 \leqslant \alpha \leqslant 2\pi$, $0 \leqslant \varepsilon \leqslant \varepsilon_0$, *tels que le difféomorphisme* $A_{\alpha,\varepsilon}$ *ne se ramène pas à une rotation par un difféomorphisme différentiable, tend vers zéro avec* ε_0.

Ce théorème est valable pour les applications d'un tore de n dimensions.

La démonstration n'entre pas dans le cadre de ce cours, cependant dans le paragraphe suivant nous verrons une technique de démonstration de tels théorèmes dans le cas élémentaire d'un difféomorphisme analytique, due à Kolmogorov.

L. Approximations de nombres irrationnels par des rationnels.

Théorème. *Pour tout nombre irrationnel* μ *il existe des approximations rationnelles aussi exactes que l'on veut, dont l'erreur est inférieure à l'inverse du carré du dénominateur*

$$\left|\mu - \frac{p}{q}\right| < \frac{1}{q^2}.$$

Le nombre π, par exemple, peut être approché avec une erreur de l'ordre du millionième par la fraction rationnelle 355/113.

Avant de prouver ce théorème on va indiquer une méthode géométrique de détermination d'une suite infinie de telles approximations (appelée algorithme des fractions continues, ou tout simplement algorithme d'Euclide).

Soit un plan rapporté aux coordonnées (x, y) (fig. 79).

Traçons une droite $y = \mu x$. Pour fixer les idées on supposera que $\mu > 0$. Repérons dans le premier quadrant tous les points à coordonnées entières. Aucun d'eux à l'exception de O n'est situé sur la droite menée, puisque μ est irrationnel. Considérons les enveloppes convexes des points entiers du quadrant, situés d'un côté (on dira « plus bas ») et de l'autre (resp. « plus haut ») de la droite. [Pour construire ces enveloppes convexes imaginons-nous un fil fixé à l'infini et superposé à la droite. Supposons qu'en chaque point entier distinct de O est planté un clou. Tirons le fil par l'extrémité libre O vers le bas (resp. vers le haut). Le fil est tendu par certains clous. La ligne polygonale ainsi obtenue constitue la frontière de l'enveloppe convexe inférieure (resp. supérieure).] Les sommets des lignes polygonales convexes construites définissent les approximations du nombre irrationnel μ dont on a parlé. Si des nombres entiers (q, p) sont les coordonnées d'un sommet, la fraction p/q correspondant à

ce sommet est dite *réduite de* μ. Pour toute réduite on a $| \mu - (p/q) | < 1/q^2$.

Pour démontrer cette inégalité nous allons décrire une autre méthode de construction de lignes polygonales convexes. Désignons par e_{-1} le vecteur de base (1, 0) et par e_0 le vecteur (0, 1). Ces vecteurs sont situés de part et d'autre de la droite $y = \mu x$. Nous allons construire une suite de vecteurs $e_1, e_2, \ldots$ d'après le principe suivant. Supposons que e_{k-1} et e_k sont déjà construits et sont situés de part

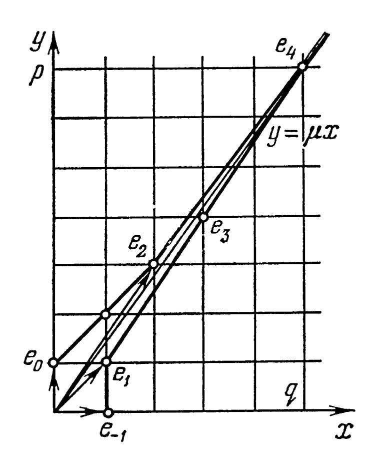


Fig. 79

et d'autre de la droite. On ajoutera au vecteur e_{k-1} le vecteur e_k autant de fois qu'il le faut pour que cette somme soit située du même côté de la droite $y = \mu x$ que le vecteur e_{k-1}.

On obtient ainsi une suite de nombres naturels a_k et la suite des vecteurs entiers

$$e_1 = e_{-1} + a_0 e_0, \ldots, e_{k+1} = e_{k-1} + a_k e_k, \ldots$$

Les vecteurs e_k sont les sommets de nos deux enveloppes convexes (alternativement supérieure pour les k pairs et inférieure pour les k impairs).

L e m m e. *L'aire du parallélogramme construit sur les vecteurs* (e_{k+1}, e_k) *est égale (compte tenu de l'orientation) à* $(-1)^k$.

◀ Pour le parallélogramme initial (e_0, e_{-1}) ceci est évident. Chaque parallélogramme possède avec celui qui le précède un côté commun et une hauteur égale, et définit l'orientation opposée du plan. ▶

C o r o l l a i r e. *Soient* q_k *et* p_k *les coordonnées du point* e_k. *La différence de deux réduites successives est alors égale à*

$$\frac{p_k}{q_k} - \frac{p_{k+1}}{q_{k+1}} = \frac{(-1)^k}{q_k q_{k+1}}.$$

◀ En réduisant ces fractions au même dénominateur, on constate que le numérateur n'est autre que le déterminant onstitué des composantes de e_{k+1} et e_k. Ce déterminant est égal à l'aire orientée du parallélogramme. ▶

Démonstration du théorème. ◀ Les vecteurs e_k sont situés alternativement soit d'un côté soit de l'autre de la droite $y = \mu x$.

Donc les réduites sont alternativement soit plus petites soit plus grandes que μ. Par suite la différence entre μ et une réduite est inférieure au module de la différence entre cette réduite et la suivante. Le corollaire nous dit que cette différence est inférieure en module à $1/q_k q_{k+1}$, ce qui est $\leqslant 1/q_k^2$, puisque $q_{k+1} \geqslant q_k$ pour $k \geqslant 0$. ▶

Remarque. Les nombres a_k sont dits *quotients incomplets*. Les réduites s'expriment en fonction des quotients incomplets comme suit:

$$\frac{p_k}{q_k} = a_0 + \cfrac{1}{a_1 + \cfrac{}{\ddots + \cfrac{1}{a_{k-1}}}}.$$

L'expression $a_0 + \cfrac{1}{a_1 + \cdots}$ s'appelle *fraction continue infinie*. Le nombre μ *se décompose en une fonction continue infinie* en ce sens que $\lim p_k/q_k = \mu$.

§ 12. Réduction analytique des difféomorphismes analytiques du cercle à une rotation

Dans ce paragraphe on se servira d'une modification de la méthode de Newton, due à Kolmogorov, pour démontrer le théorème des difféomorphismes analytiques du cercle proches d'une rotation et possédant pour nombre de rotation un nombre presque arbitraire.

A. Enoncé du théorème.

Soient Π_ρ la bande $|\operatorname{Im} y| < \rho$, a une fonction holomorphe bornée sur cette bande, $\|a\|_\rho = \sup |a(y)|$, $y \in \Pi_\rho$.

Soient μ un nombre irrationnel, $K > 0$, $\sigma > 0$. On dira que μ est un *nombre de type* (K, σ) si, quels que soient p et $q \neq 0$, on a

$$|\mu - (p/q)| \geqslant K/|q|^{2+\sigma}.$$

Théorème. *Il existe un* $\varepsilon > 0$ *dépendant uniquement de* K, ρ *et* σ, *tel que si* a *est une fonction* 2π*-périodique analytique et réelle sur l'axe réel telle que* $\|a\|_\rho < \varepsilon$ *et que l'application*

$$y \mapsto y + 2\pi\mu + a(y)$$

définisse un difféomorphisme du cercle dont le nombre de rotation μ *est du type* (K, σ), *alors ce difféomorphisme est analytiquement équivalent à une rotation d'angle* $2\pi\mu$.

B. Equation homologique.

Désignons par $\mathfrak{U}$ une rotation d'angle $2\pi\mu$ et par H le difféomorphisme cherché transformant cette rotation en A : le diagramme suivant est commutatif

$$\begin{array}{ccc} S^1 & \xrightarrow{A} & S^1 \\ H\uparrow & & \uparrow H \\ S^1 & \xrightarrow{\mathfrak{U}} & S^1 \end{array} \quad \text{i.e.} \quad H \circ \mathfrak{U} = A \circ H.$$

Mettons H sous la forme $Hz = z + h(z)$, $h(z + 2\pi) = h(z)$. On obtient alors l'équation fonctionnelle en h

$$h(z + 2\pi\mu) - h(z) = a(z + h(z)).$$

i A diffère peu d'une rotation, alors a est petit. Il semble naturel que h soit du même ordre de petitesse. Alors $a(z + h(z))$ diffère de $a(z)$ d'une quantité d'ordre de petitesse supérieur à a. Donc en « première approximation » on obtient pour h l'équation

$$h(z + 2\pi\mu) - h(z) = a(z).$$

Cette équation linéaire s'appelle *équation homologique.*

Remarque. Nous pouvons considérer l'ensemble de tous les difféomorphismes A comme une « variété de dimension infinie » sur laquelle opère un « groupe de dimension infinie » de difféomorphismes H. La fonction a peut être interprétée comme le vecteur tangent à la variété de difféomorphismes au point $\mathfrak{U}$, et la fonction h comme le vecteur tangent au groupe en l'unité.

Dans ces termes l'équation homologique admet l'interprétation suivante : a appartient à l'espace tangent à l'orbite du point $\mathfrak{U}$ sous l'action du groupe si et seulement si l'équation homologique en h est résoluble.

C. Solution formelle de l'équation homologique.

Développons la fonction a et la fonction inconnue h en séries de Fourier

$$a(z) = \sum a_k e^{ikz}, \quad h = \sum h_k e^{ikz}.$$

L identification des coefficients en e^{ikz} donne

$$h_k = \frac{a_k}{e^{2\pi ik\mu} - 1}.$$

Pour que cette équation admette une solution il est nécessaire que les dénominateurs s'annulent en même temps que les numérateurs. En particulier, l'équation homologique est insoluble si $a_0 \neq 0$. Si $a_0 = 0$ et si le nombre de rotation μ est irrationnel, alors les formules précédentes définissent la solution de l'équation homologique dans la classe des séries formelles de Fourier. Pour obtenir la olution véritable il faut étudier la convergence de cette série.

D. Comportement des coefficients de Fourier de fonctions analytiques.

Lemme 1. *Si f est une fonction 2π-périodique, analytique dans la bande Π_ρ, continue dans l'adhérence de cette bande et $\| f \|_\rho \leqslant M$, alors les coefficients de Fourier décroissent comme une progression géométrique*

$$|f_k| \leqslant M e^{-|k|\rho}.$$

◀ On sait que $f_k = \frac{1}{2\pi} \oint f(z) e^{-ikz} dz$. Soit $k > 0$. Déplaçons le chemin d'intégration vers le bas (de $-i\rho$). Cette intégrale ne varie pas, puisque les inté-

grales étendues aux côtés verticaux du rectangle obtenu sont égales. Donc

$$f_k=\frac{1}{2\pi}\int_0^{2\pi} f(x-i\rho)\, e^{-ikx-k\rho}\, dx, \quad |f_k| \leqslant Me^{-k\rho}.$$

Pour $k<0$ il faut déplacer le chemin d'intégration vers le haut (de $i\rho$). ▶

Lemme 2. *Si* $|f_k| \leqslant Me^{-|k|\rho}$, *la fonction* $f=\sum f_k e^{ikz}$ *est analytique dans la bande* Π_ρ *et, de plus,*

$$\| f \|_{\rho-\delta} \leqslant 4M/\delta \text{ pour } \delta<\rho,\ \delta<1.$$

$$◀ \| f \|_{\rho-\delta} \leqslant \sum |f_k|\,|e^{ikz}| \leqslant M\sum e^{-|k|\rho} e^{|k|(\rho-\delta)} =$$

$$=M\sum e^{-|k|\delta} \leqslant 2M/(1-e^{-\delta}) \leqslant 4M/\delta. ▶$$

Remarque. Dans le cas de fonctions à n variables le lemme 1 reste valable tandis que dans le lemme 2 la majoration $4M/\delta$ est remplacée par CM/δ^n, où $C=C(n)$ est une constante indépendante de δ et de f.

E. Petits diviseurs.

Pour résoudre l'équation homologique il faut diviser les coefficients de Fourier du second membre par les nombres $e^{2\pi ik\mu}-1$. Si le nombre μ est irrationnel, ces nombres sont non nuls pour $k\neq 0$. Cependant certains d'entre eux sont très proches de zéro. En effet, tout nombre μ admet des approximations rationnelles p/q avec une erreur de $|\mu-(p/q)|<1/q^2$ pour q aussi grand que l'on veut. Pour $k=q$ le diviseur $e^{2\pi ik\mu}-1$ sera très petit.

On établit que ces petits diviseurs sont minorés presque sûrement par une puissance de k.

Lemme 3. *Soit* $\sigma>0$. *Alors pour presque chaque réel* μ *il existe* $K=K(\mu,\sigma)>0$ *tel que*

$$\left|\mu-\frac{p}{q}\right| \geqslant \frac{K}{|q|^{2+\sigma}}$$

pour les p *et* q *entiers non nuls.*

◀ Considérons sur l'intervalle $[0, 1]$ les nombres μ pour lesquels l'inégalité précédente (avec p, q, K, σ fixes) est violée. Ces nombres forment un intervalle de longueur $\leqslant 2K/q^{2+\sigma}$. La réunion de ces intervalles pour tous les p ($q>0$, K et σ étant fixes) est de longueur $\leqslant 2K/q^{1+\sigma}$. En sommant sur q on obtient un ensemble de mesure $\leqslant CK$, où $C=2\sum q^{-(1+\sigma)}<\infty$. Donc l'ensemble des nombres $\mu\in[0, 1]$. pour lesquels n'existe pas la quantité K du lemme, est recouvert par des ensembles de mesure aussi petite que l'on veut. Donc cet ensemble est de mesure nulle (sur l'intervalle $[0, 1]$ et, par conséquent, sur toute la droite).

Remarque. Les nombres μ réalisant l'inégalité précédente sont appelés au point A nombres de type (K, σ).

Pour un nombre μ *de type* (K, σ) *le petit diviseur admet la minoration suivante*

$$|e^{2\pi ik\mu}-1| \geqslant K/(2|k|^{1+\sigma}) \quad (|k|>0).$$

◀ En effet, la distance de $k\mu$ au plus proche entier est minorée par le nombre $K/|k|^{1+\sigma}$, quant à toute corde du cercle unité, elle n'est pas plus courte que la longueur du plus petit des arcs qu'elle intercepte, divisée par π. ▶

F. Etude de l'équation homologique.

Soit a une fonction analytique 2π-périodique de valeur moyenne zéro.

L e m m e 4. *Pour presque tous les* μ *l'équation homologique possède une solution analytique* 2π*-périodique (réelle si a l'est). Il existe une constante* $\nu = \nu(K, \sigma) > 0$ *telle que si* μ *est de type* (K, σ), *alors pour tout* $\delta > 0$, *inférieur à* ρ *et pour* $\rho < 1/2$ *on a*

$$\| h \|_{\rho-\delta} \leqslant \| a \|_\rho \, \delta^{-\nu}.$$

R e m a r q u e. Le passage de a à h n'altère pas plus les propriétés de la fonction qu'une ν-uple dérivation. [Il est utile de remarquer que $\| d^\nu f/dz^\nu \|_{\rho-\delta} \leqslant C \| f \|_\rho \delta^{-\nu}$ en vertu de la majoration de Cauchy des coefficients du développement taylorien.] Si l'on néglige l'altération de la fonction due à une ν-uple dérivation, on peut dire que la solution h de l'équation homologique est du même ordre de petitesse que son second membre a.

◀ 1°. Le lemme 1 dit que $| a_k | \leqslant M e^{-|k|\rho}$ si $\| a \|_\rho \leqslant M$.

2°. μ étant de type (K, σ), on a $| h_k | \leqslant 2M e^{-|k|\rho} | k |^{1+\sigma}/K$.

3°. La fonction $x^m e^{-\alpha x}$, $x \geqslant 0$ possède un maximum au point $x = m/\alpha$. Donc $x^m e^{-\alpha x} \leqslant C\alpha^{-m}$, $C = (m/e)^m$, $\forall \alpha > 0$, $\forall x > 0$. Donc $\forall \alpha > 0$

$$| k |^{1+\sigma} e^{-\alpha | k |} \leqslant C\alpha^{-m}, \quad m = 1 + \sigma.$$

4°. Ainsi, $|h_k| \leqslant M e^{-|k|(\rho-\alpha)} 2CK^{-1}\alpha^{-m}$. D'après le lemme 2 on a $\| h \|_{\rho-\delta} \leqslant DM$, où $D = 8C/K\alpha^m (\delta - \alpha)$. Prenons $\alpha = \delta/2$. Le nombre $D \leqslant \delta^{-\nu}$ si ν est assez grand (car $< 1/2$). ▶

G. Construction des approximations successives.

Résolvons l'équation homologique de second membre $\tilde{a} = a - a_0$ (a_0 es la valeur moyenne de la fonction a). Soit h^0 une solution. Soit l'application $H_0 : z \to z + h^0(z)$. Construisons l'application $A_1 = H_0^{-1} \circ A \circ H_0$. Soit la fonction a^1 définie par la relation $A_1 z = z + 2\pi\mu + a^1(z)$.

En d'autres termes, nous avons introduit sur le cercle une nouvelle coordonnée z_1 (où $z = H_0(z_1)$) et exprimé l'application A dans cette nouvelle coordonnée. L'application $z_1 \longmapsto A_1 z_1$ obtenue diffère d'une rotation d'angle $2\pi\mu$ d'un « résidu » a^1.

L'approximation suivante se construit de la même façon à partir de A_1 au lieu de A. On construit h^1 et le changement H_1 qui envoie A_1 dans $A_2 = H_1^{-1} \circ A_1 \circ H_1$.

On obtient une suite de changements H_n. Considérons le changement $\mathscr{H}_n = H_0 \circ H_1 \circ \ldots \circ H_{n-1}$. On a $A_n = \mathscr{H}_n^{-1} \circ A \circ \mathscr{H}_n$.

La suite $\mathscr{H}_n$ *est convergente si* μ *est de type* (K, σ) *et si* $\| a \|_\rho$ *est assez petit.* Le changement limite $\mathscr{H}$ envoie l'application initiale des $\mathscr{H}^{-1} \circ A \circ \mathscr{H} = \lim A_n$ qui est une rotation d'angle $2\pi\mu$.

H. Estimation du résidu après une approximation.

L e m m e 5. *Il existe des constantes* $\varkappa$, $\lambda > 0$ *dépendant uniquement de* K *et* σ, *telles que* $\forall \delta \in]0, \rho[$, $\rho < \frac{1}{2}$, on ait

$$\| a \|_\rho \leqslant \delta^\varkappa \Rightarrow \| a^1 \|_{\rho-\delta} \leqslant \| a \|_\rho^2 \, \delta^{-\lambda}.$$

R e m a r q u e. Ceci signifie que le résidu a^1 qui subsiste après le premier changement de variables est du deuxième ordre de petitesse par rapport au résidu initial a de la rotation (à l'altération près du type λ-uple dérivation de la fonction). Donc, dans le schéma indiqué des approximations successives l'erreur commise sur toute approximation est de l'ordre du carré de l'erreur commise sur la précédente. Au terme de n approximations on obtient une erreur de l'ordre de ε^{2^n}, où ε est l'erreur affectant l'approximation initiale.

Cette convergence caractéristique de la méthode des tangentes de Newton (fig. 80) permet d'annihiler l'influence des petits diviseurs qui apparaissent à chaque pas (c'est-à-dire l'influence du facteur altérant $\delta^{-\lambda}$). Cette méthode de lutte contre les petits diviseurs est due à A. Kolmogorov (1954).

◀ 1°. Supposons que Ω est un domaine convexe de $\mathbb{C}^n$ (ou $\mathbb{R}^n$), $h: \Omega \to \mathbb{C}^n$ (resp. $\mathbb{R}^n$) une application différentiable telle que $\| h_* \| = \sup_{x \in \Omega} \| h_*(x) \| < 1$. *L'application H qui envoie x dans $x + h(x)$ est un difféomorphisme de Ω sur $H\Omega$.*

◀ Les valeurs propres de $H_*(x)$ sont différentes de zéro, donc H est un difféomorphisme local. L'application h est contractante en vertu de la condition $| h_* | < q < 1$ et de la convexité de Ω. Donc la différence entre les images

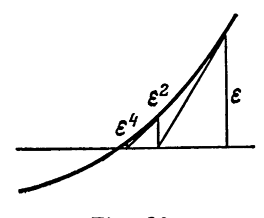


Fig. 80

de deux points quelconques par l'application H est inférieure à la distance de ces points, donc leurs images sont distinctes, c'est-à-dire H est une bijection.▶

2°. Montrons que *si $\varkappa$ est assez grand l'application A_1 est analytique dans la bande $\Pi_{\rho-\delta}$.*

◀ Soit $\| a \|_\rho \leqslant M = \delta^\varkappa$. Alors $| a_0 | \leqslant M$, $\| \tilde{a} \|_\rho \leqslant 2M$. Le théorème du point D nous dit que $\| h^0 \|_{\rho-\alpha} \leqslant 2M\alpha^{-\nu}$. Donc $\| dh^0/dz \|_{\rho-2\alpha} \leqslant 2M\alpha^{-(\nu+1)}$.

Prenons $\alpha = \delta/8$. Si $\varkappa$ est assez grand les inégalités précédentes donnent $\| a \|_\rho < \alpha$, $\| h^0 \|_{\rho-\alpha} < \alpha$, $\| dh^0/dz \|_{\rho-2\alpha} < \alpha$.

Donc en vertu du 1°, H_1 est un difféomorphisme de la bande $\Pi_{\rho-2\alpha}$ et l'image contient la bande $\Pi_{\rho-3\alpha}$.

On a maintenant $H_0\Pi_{\rho-\delta} \subset \Pi_{\rho-\delta+\alpha}$, $A \circ H_0\Pi_{\rho-\delta} \subset \Pi_{\rho-\delta+2\alpha} \subset \Pi_{\rho-3\alpha}$. Donc le difféomorphisme H_0^{-1} est défini sur $A \circ H_0\Pi_{\rho-\delta}$. Par suite, l'application $A_1 = H_0^{-1} \circ A \circ H_0$ est analytique dans $\Pi_{\rho-\delta}$ et est un difféomorphisme. ▶

3°. *Evaluons le résidu* a^1. ◀ Le diagramme commutatif définissant a^1 donne

$$z + 2\pi\mu + a^1(z) + h^0(z + 2\pi\mu + a^1(z)) \equiv z + h^0(z) + 2\pi\mu + a(z + h^0(z)).$$

Compte tenu de l'équation homologique on a

$$a^1(z) = [a(z + h^0(z)) - a(z)] - [h^0(z + 2\pi\mu + a^1(z)) - h^0(z + 2\pi\mu)] + a_0.$$

Le premier crochet s'évalue à l'aide du théorème de la moyenne et de l'inégalité de Cauchy. En vertu du 2°, on a

$$\| a(z + h^0(z)) - a(z) \|_{\rho-\delta} \leqslant \frac{M}{\delta} \| h^0 \|_{\rho-\delta} \leqslant M^2\delta^{-u},$$

où la constante u dépend uniquement de ν, c'est-à-dire uniquement de K et σ.

Le deuxième crochet s'estime de façon analogue:

$$\| [\ ] \|_{\rho-\delta} \leqslant 2M\alpha^{-(\nu+1)} \| a^1 \|_{\rho-\delta} \leqslant M\delta^{-u_1} \| a^1 \|_{\rho-\delta}.$$

Donc

$$\| a^1 \|_{\rho-\delta} (1 - M\delta^{-u_1}) \leqslant | a_0 | + M^2\delta^{-u}. \blacktriangleright$$

4°. *Evaluons maintenant la quantité* $| a_0 |$ en utilisant le fait que le nombre de rotation de l'application A, et, partant de l'application A_1, est égal à $2\pi\mu$.

◀ Il résulte que a^1 s'annule en un point réel z_0. Portons z_0 dans la formule de $a^1(z)$. On obtient $a_0 = a(z_0) - a(z_0 + h^0(z_0))$ et par suite $| a_0 | \leqslant M^2\delta^{-u}$ (cf. 3°). ▶

5°. Les majorations 3° et 4° entraînent $\| a^1 \|_{\rho-\delta} \leqslant 4M^2\delta^{-u}$. ▶

I. Convergence du système d'approximations.

1°. Etudions l'application A_n construite au n-ème pas dans la bande dont le rayon ρ_n décroît à chaque approximation: $\rho_0 = \rho$, $\rho_n = \rho_{n-1} - \delta_{n-1}$. Nous choisirons la suite de nombres δ_n décroissante et telle que

$$\delta_n = \delta_{n-1}^{3/2}, \; \delta_0 < 1/2.$$

Pour δ_0 assez petit on aura

$$\sum \delta_n < \rho/2.$$

2°. Formons une suite de nombres M_n en posant

$$M_n = \delta_n^N.$$

Le nombre N (qui est assez grand et dépend uniquement de K et σ) sera définitivement choisi plus bas. Pour l'instant on remarquera que $M_n = M_{n-1}^{3/2}$.

3°. *Supposons que* $\| a \|_\rho \leqslant M_0$. *Montrons que* $\| a^n \|_{\rho_n} \leqslant M_n$.

◀ La proposition 3 nous dit que si $N > \varkappa$, alors

$$\| a^1 \|_{\rho_1} \leqslant M_0^2 \delta_0^{-1} = \delta_0^{2N-\lambda}.$$

Or $\delta_0^{2N-\lambda} < \delta_1^N = \delta_0^{3N/2}$ si $N > 2\lambda$. Prenons N plus grand que 2λ et $\varkappa$. On aura

$$\| a^1 \|_{\rho_1} \leqslant \delta_1^N = M_1.$$

On passe de façon analogue de a^{n-1} à a^n. ▶

4°. *Prouvons la convergence des produits* $\mathcal{H}_n = H_0 \circ \ldots \circ H_{n-1}$ *dans* $\Pi_{\rho/2}$.

Le difféomorphisme H_0 est analytique dans Π_{ρ_1} et vérifie les inégalités. $\| h^0 \|_{\rho_1} \leqslant \delta_0$, $\| dh^0/dz \|_{\rho_1} \leqslant \delta_0$ (cf. 2° point H).

De façon analogue pour H_{n-1} on obtient

$$\| h^{n-1} \|_{\rho_n} \leqslant \delta_{n-1}, \quad \| dh^{n-1}/dz \|_{\rho_n} \leqslant \delta_{n-1}.$$

Donc $\mathcal{H}_n$ est analytique dans Π_{ρ_n} et possède une dérivée bornée supérieurement et inférieurement par les quantités $C = \prod (1 + \delta_k)$, $c = \prod (1 - \delta_k)$.

D'où il suit que $\mathcal{H}_n$ est un difféomorphisme de Π_{ρ_n} et que la suite $\mathcal{H}_n$ est convergente dans $\Pi_{\rho/2}$. En effet,

$$\| \mathcal{H}_n - \mathcal{H}_{n+1} \|_{\rho/2} \leqslant C \| h^n \|_{\rho/2} \leqslant C\delta_n.$$

Soit H la limite de la suite $\mathcal{H}_n$. En passant à la limite dans la relation $A \circ \mathcal{H}_n = \mathcal{H}_n \circ A_n$ on obtient $A \circ \mathcal{H} = \mathcal{H} \circ \mathfrak{U}$, où $\mathfrak{U}$ est une rotation d'angle $2\pi\mu$. Le théorème est démontré.

J. Remarques.

1°. J. Moser s'est aperçu qu'en combinant les approximations décrites avec le lissage de J. Nash on peut prouver un théorème analogue dans le cas d'une différentiabilité finie (J. Moser, A rapidly convergent iteration method and nonlinear partial differential equations. I, Ann. Scuola Norm Pisa 3) 20; 265-316; II 20: 499).

Les premiers travaux de Moser ont nécessité des centaines de dérivées. Par la suite Moser et Russmann ont réussi à réduire ce nombre (H. Russmann, Kleine Nenner II: Bemerkungen zur Newtonischen Methode, Nachr. Acad. Wiss. Göttingen, Math. Phys. Klasse 1 (1972), 1-10).

2°. Dans le cas multidimensionnel le nombre de rotation n'est pas défini. Cependant dans la famille d'applications $y \mapsto y + \alpha + a(y)$, a étant petit, $y \in T^n$, pour la plupart des α l'application est différentiablement équivalente à la translation $y \mapsto y + 2\pi\mu$. En particulier, pour la famille analytique $y \mapsto$ $\mapsto y + \alpha + \varepsilon a_1(y) + \varepsilon^2 a_2(y) + \ldots$, pour presque tout μ il existe une fonction analytique $\alpha(\varepsilon) = 2\pi\mu + \varepsilon\mu_1 + \ldots$ telle que l'application $y \mapsto y +$ $+ \alpha(\varepsilon) + \varepsilon a_1(y) + \ldots$ se transforme en $y \mapsto y + 2\pi\mu$ par le changement analytique $y = z + \varepsilon h_1(z) + \ldots$.

On peut déterminer les coefficients $h_1, \ldots$ en identifiant les termes de même puissance Mais on ne démontre la convergence des séries ainsi obtenues qu'indirectement à l'aide des approximations de Newton.

3°. Il semble qu'un difféomorphisme analytique du cercle est analytiquement équivalent à une rotation irrationnelle si et seulement si l'ensemble des points fixes des puissances de ce difféomorphisme ne possède pas de points limites réels. On peut encore penser que pour certains μ irrationnels anormalement bien approchés par des nombres rationnels, la fonction $\alpha(\varepsilon)$ décrite au 2° n'est même pas indéfiniment différentiable (même dans le cas unidimensionnel).

§ 13. Introduction à la théorie hyperbolique

Dans ce paragraphe on démontre le théorème d'Anossov de la stabilité structurelle d'un automorphisme du tore et le théorème de Grobman-Hartman de la stabilité structurelle du col.

A. Exemple élémentaire: automorphisme linéaire du tore.

Les équations différentielles à espaces des phases multidimensionnels définissent une riche classe de systèmes structurellement stables

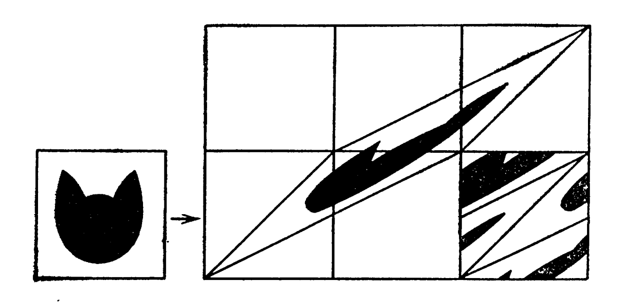


Fig. 81

dans lesquels toute trajectoire est disposée parmi les autres comme une position d'équilibre du type col parmi les hyperboles voisines. Commençons par un exemple élémentaire (fig. 81).

Soit A automorphisme du tore T^2, défini par $\hat{A}$, application linéaire entière unimodulaire (de déterminant 1) du plan, de matrice

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Le réseau $2\pi\mathbb{Z}^2$ est transformé dans lui-même par l'application $\hat{A}$. Les images des points équivalents (congruents modulo 2π) du plan par l'application $\hat{A}$ sont des points équivalents. Donc $\hat{A}$ définit une application A du tore sur lui-même. La matrice inverse $\hat{A}^{-1}$ est aussi entière puisque $\det \hat{A} = 1$. Donc A est un difféomorphisme du tore sur lui-même. De plus, A est un automorphisme du groupe $T^2 = \mathbb{R}^2/2\pi\mathbb{Z}^2$.

B. Propriétés de l'automorphisme du tore.

Un ensemble fini de points est un cycle d'une application A si A les permute en cycle.

Théorème 1. *L'automorphisme A du tore possède un nombre dénombrable de cycles. Tous les points dont les deux coordonnées sont des multiples rationnels de 2π et eux seuls sont points des cycles de A.*

◀ 1°. Figeons un entier N; les points du tore dont les coordonnées sont des multiples entiers de 2π divisés par N forment un ensemble fini. L'application A envoie cet ensemble dans lui-même. Par suite, tous les points de cet ensemble appartiennent à des cycles.

2°. Soit $2\pi\xi$ un point d'un cycle d'ordre $n > 1$. Alors $\hat{A}^n\xi = \xi + m$, où m est un vecteur entier. L'équation linéaire obtenue en ξ possède un dénominateur non nul. Donc les composantes de ξ sont rationnelles. ▶

Théorème 2. *Les itérations de l'automorphisme A étalent un domaine quelconque F uniformément sur le tore: pour tout domaine G*

$$\lim_{n\to\infty} \frac{\operatorname{mes}(A^nF) \cap G}{\operatorname{mes} F} = \frac{\operatorname{mes} G}{\operatorname{mes} T^2}.$$

Cette propriété de l'automorphisme A est dite *brassage* (ou « mixing »); elle a lieu quels que soient les ensembles mesurables F et G.

◀ En termes de fonctions sur le tore, cette relation s'écrit

$$\lim_{n\to\infty} (A^{n*}f, g) = (f, 1)(1, g)/(1, 1),$$

où $(u, v) = \int u(x)\bar{v}(x)\,dx$, $(A^{n*}f)(x) = f(A^nx)$.

Soit f exponentielle: $f = e^{i(p,x)}$. Alors $A^{n*}f$ est aussi exponentielle de vecteur d'onde $p' = \hat{A}^{n\prime}p$. Si $p \neq 0$ l'orbite du point p est infinie sous l'action des opérateurs $\hat{A}^{n\prime}$. Donc pour toute exponentielle $g = e^{i(q,x)}$ on a $\lim_{n\to\infty} (A^{n*}f, g) = 0$. En approchant f et g en moyenne quadratique par des sommes d'exponentielles on obtient ce qu'on voulait. ▶

Voici une démonstration du brassage plus instructive (quoique plus complexe si elle est effectuée soigneusement).

Théorème 3. *Il existe sur le tore deux champs de directions invariants par l'automorphisme A. Les courbes intégrales de chacun d'eux sont partout denses sur le tore. L'automorphisme A envoie courbes intégrales dans courbes intégrales du même champ, en les dilatant de λ fois* ($\lambda > 1$) *dans un champ et les contractant de λ fois dans l'autre* (fig. 82).

◀ Considérons les valeurs propres $\lambda_{1,2} = \frac{3 \pm \sqrt{5}}{2}$ de l'application $\hat{A}$. De toute évidence, $\lambda_1 > 1 > \lambda_2$ et λ_1 et λ_2 sont irrationnels.

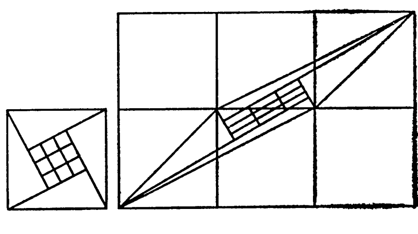


Fig. 82

Considérons sur le plan la famille de toutes les droites parallèles au premier vecteur propre de l'application $\hat{A}$. Du fait que λ_1 est irrationnel, les composantes du vecteur propre sont incommensurables. Par suite, les droites de la famille définissent sur le tore des hélices partout denses. L'application $\hat{A}$ envoie chaque droite dans une droite de la même famille, en la dilatant de $\lambda_1 > 1$ fois. Par suite, l'application A envoie les hélices dans des hélices de la même famille en les dilatant de λ_1 fois. Cette famille est dite *feuilletage dilatant* de l'application A.

La deuxième direction propre définit respectivement un *feuilletage contractant.*

Considérons maintenant l'image d'un domaine plan $\hat{F}$ par l'application A^n du plan. Cette application est une rotation hyperbolique : le produit d'une dilatation de rapport λ_1^n suivant la première direction propre par une contraction de rapport λ_1^n suivant la seconde. Donc l'image du domaine $\hat{F}$ pour n grand est une bande longue, étroite, s'étirant le long de la première direction propre. Par suite, l'image de $\hat{F}$ sur le tore par l'application A^n est une bande longue, étroite, proche d'une portion de trajectoire de l'équation $\dot{x} = \omega$, ω étant un vecteur non résonnant. Or cette trajectoire est uniformément distribuée sur le tore. Donc lorsque n croît les images $A^n\check{F}$ coupent tout domaine G du tore ; en poursuivant nos raisonnements on établit la propriété de brassage.

C. Stabilité structurelle de l'automorphisme du tore.

L'automorphisme du tore étudié plus haut est structurellement stable dans la classe des difféomorphismes du tore; ce résultat étonnant établi au début des années 60 constitue depuis lors l'un des faits les plus remarquables de la théorie des équations différentielles. En particulier, tout difféomorphisme assez proche de A possède un nombre dénombrable de cycles et un ensemble partout dense de points périodiques.

Théorème d'Anossov. *L'automorphisme du tore* $A : T^2 \to T^2$ *de matrice* $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ *est structurellement stable pour la topologie* C^1. *En d'autres termes, tout difféomorphisme B assez proche de A avec sa dérivée est conjugué de A par un homéomorphisme H : $B = H^{-1} \circ A \circ H$.*

Remarque. L'homéomorphisme H peut être choisi aussi proche que l'on veut de l'identité si B est assez proche de A, mais généralement ne peut être rendu différentiable.

Le théorème d'Anossov indique que dans le cas de systèmes à espaces des phases multidimensionnels il est possible que de faibles déformations n'influent pas sur le comportement des trajectoires ne se rapprochant pas d'une position d'équilibre stable ou de cycles, comme c'était le cas pour les champs de vecteurs sur une sphère à deux dimensions ou un tore. Nous étudierons plus bas l'interprétation physique de ce comportement des systèmes dynamiques, qui est plus complexe que les auto-oscillations. La démonstration du théorème d'Anossov est donnée dans les points D à G.

D. Equation homologique.

On cherche l'homéomorphisme H: $H(x) = x + h(x)$ rendant commutatif le diagramme

$$\begin{array}{ccc} \mathbb{R}^2 & \xrightarrow{B} & \mathbb{R}^2 \\ \uparrow & & \uparrow H, \\ \mathbb{R}^2 & \xrightarrow{A} & \mathbb{R}^2 \end{array}$$

où $B(x) = Ax + f(x)$; les fonctions f et h sont 2π-périodiques en x.

A partir du diagramme on déduit l'équation fonctionnelle non linéaire en h:

$$h(Ax) - Ah(x) = f(x + h(x)).$$

On admet que la fonction f est petite et que h est du même ordre de petitesse; on peut donc au second membre négliger les termes du second ordre de petitesse et conserver seulement $f(x)$. On obtient l'équation linéarisée

$$h(Ax) - Ah(x) = f(x)$$

appelée *équation homologique.*

E. Résolution de l'équation homologique.

Le premier membre de l'équation homologique dépend linéairement de h. Soit L l'opérateur linéaire associant à h le premier membre de l'équation homologique. La solution s'écrit $h = L^{-1}f$. Il suffit de démontrer que l'opérateur L est inversible.

L e m m e 1. *L'espace des champs de vecteurs sur le tore se décompose en une somme directe de deux sous-espaces invariants par L.*

◀ Les espaces des champs de vecteurs parallèles à la première et la seconde direction propre de l'opérateur A sont invariants par l'opérateur L et tout champ de vecteurs est représenté de façon unique par une somme de deux champs parallèles aux directions propres de A. ▶

Soient $f = f_1 e_1 + f_2 e_2$, $h = h_1 e_1 + h_2 e_2$ les décompositions des champs f et h. L'équation homologique se transforme alors en le système

$$h_1(Ax) - \lambda_1 h_1(x) = f_1(x),$$
$$h_2(Ax) - \lambda_2 h_2(x) = f_2(x),$$

où $\lambda_1 = \lambda_2^{-1} > 1 > \lambda_2 = \lambda$ sont les valeurs propres.

Soit S l'opérateur de A-translation de l'argument dans l'espace des fonctions continues sur le tore. On a

$$(Sg)(x) = g(Ax), \quad \| S \| = 1, \quad \| S^{-1} \| = 1.$$

L'équation homologique s'écrit maintenant

$$(S - \lambda_i E) h_i = f_i, \quad i = 1, 2.$$

Soit $i = 1$. On a alors

$$(S - \lambda_1 E)^{-1} = -\lambda (E + \lambda S + \lambda^2 S^2 + \ldots).$$

Du fait que $\lambda < 1$ et $\| S \| = 1$, l'opérateur inverse existe et

$$\| (S - \lambda_1 E)^{-1} \| \leqslant \lambda/(1 - \lambda).$$

De façon analogue,

$$- \lambda_2 E)^{-1} = S^{-1}(E - \lambda S^{-1})^{-1} = S^{-1}(E + \lambda S^{-1} + \lambda^2 S^{-2} + \ldots),$$
$$\| (S - \lambda_2 E)^{-1} \| \leqslant 1/(1 - \lambda).$$

Donc, l'opérateur L^{-1} existe et, de plus, $\| L^{-1} \| \leqslant 1/(1 - \lambda)$. L'équation homologique est résolue.

F. Construction de l'application H.

L'équation fonctionnelle non linéaire du point D se résout maintenant par la méthode élémentaire des applications contractantes. Posons

$$\Phi[h](x) = f(x + h(x)) - f(x).$$

L'équation fonctionnelle s'écrit

$$Lh = \Phi h + f, \quad h = L^{-1}\Phi h + L^{-1}f.$$

L e m m e 2. *Si la norme de f dans C^1 est assez petite, alors l'opérateur $L^{-1}\Phi$ est contractant dans l'espace C^0.*

◀ Il suffit de vérifier que l'opérateur non linéaire réalise la condition de Lipchitz dans un petit rapport. En effet, en vertu du point D on a

$$\| L^{-1}\Phi h^1 - L^{-1}\Phi h^2 \| \leqslant \| \Phi h^1 - \Phi h^2 \|/(1 - \lambda).$$

Or

$$\| \Phi h^1 - \Phi h^2 \| = \max | f(x + h^1(x)) - f(x + h^2(x)) | \leqslant$$
$$\leqslant \| f \|_{C^1} \| h^1 - h^2 \|.$$

Donc l'opérateur $L^{-1}\Phi$ est contractant si $\| f \|_{C^1} < 1 - \lambda$. ▶ Si cette condition est réalisée l'équation non linéaire est résolue et H construite.

G. Propriétés de l'application H.

Prouvons que H est un homéomorphisme du tore.

◀ Si h est petit dans la métrique de C^1, l'application $H = E + h$ est un homéomorphisme. Nous savons seulement que h est petit dans la métrique de C^0. Cependant en vertu des propriétés hyperboliques de l'application $\hat{A}$ du plan, il résulte de $H(x) = H(y)$ que $x = y$.

En effet, $\hat{B}\hat{H} = \hat{H}\hat{A}$ sur le plan. Donc $\hat{H}\hat{A}\hat{x} = \hat{H}\hat{A}\hat{y}$ et, d'une façon générale, $\hat{H}\hat{A}^n\hat{x} = \hat{H}\hat{A}^n\hat{y}$. Or l'hyperbolicité de $\hat{A}$ entraîne que la distance entre les points $\hat{A}^n\hat{x}$ et $\hat{A}^n\hat{y}$ tend vers l'infini soit pour $n \to \infty$, soit pour $n \to -\infty$. Ceci contredit le fait que h est bornée. Par suite, $\hat{x} = \hat{y}$ et $x = y$ sur le tore.

Montrons que l'image de l'application H est le tore tout entier. En effet, l'image d'un cercle de rayon assez grand sur le plan par $\hat{H}$ contient le cercle de rayon 2π (puisque h est bornée), donc $HT^2 = T^2$. Par suite, H est un homéomorphisme du tore et, de plus, $BH = HA$. ▶

Le théorème du point C est prouvé.

H. Théorème de la stabilité structurelle du col.

Les raisonnements précédents démontrent également la proposition suivante.

T h é o r è m e d e G r o b m a n - H a r t m a n. *Soit: $A : \mathbb{R}^n \to \mathbb{R}^n$ une application linéaire ne possédant pas de valeurs propres égales à l'unité en module. Tout difféomorphisme local $B : (\mathbb{R}^n, O) \to (\mathbb{R}^n, O)$ de partie linéaire A en un point fixe O est topologiquemen équivalent à A dans un voisinage assez petit de O.*

◀ Au voisinage du point O le difféomorphisme local B coïncide avec le difféomorphisme global $C : \mathbb{R}^n \to \mathbb{R}^n$ défini comme suit:

soit φ une fonction différentiable égale à zéro en dehors d'un 1-voisinage du point O et à l'unité dans un petit voisinage du point O. Alors C coïncide avec A en dehors d'un ε-voisinage dans lequel φ_ε est différent de O et égal à $A + \varphi_\varepsilon (B - A)$ à l'intérieur de ce voisinage ($\varphi_\varepsilon (x) = \varphi (x/\varepsilon)$).

La précédente démonstration du théorème nous dit que tout difféomorphisme $\mathbb{R}^n \to \mathbb{R}^n$, C^1-proche de A, est topologiquement équivalent à A. Or la différence $C - A$ peut être rendue C^1-petite par un choix convenable de $\varepsilon > 0$, puisque dans un ε-voisinage de zéro on a

$$| B - A | \leqslant c\varepsilon^2, \quad | (B - A)' | \leqslant c\varepsilon.$$

Par suite, C est topologiquement équivalent à A.

Or C comme A ne possède qu'un point fixe O. Donc l'homéomorphisme qui envoie A dans C laisse le point O fixe. ▶

§ 14. C-systèmes

Dans ce paragraphe on définit les C-difféomorphismes et les C-flots et on se penche sur leurs applications en théorie des flots géodésiques sur des variétés de courbure négative et dans d'autres problèmes.

A. Définition d'un C-difféomorphisme.

L'analyse de l'automorphisme du tore envisagé plus haut montre que seuls les feuilletages contractants et dilatants sont essentiels dans les raisonnements précédents; on peut donc donner une définition générale du difféomorphisme hyperbolique sans supposer que M est un tore.

Soit $A: M \to M$ difféomorphisme d'une variété compacte. Supposons que

1) tout espace tangent à M est la somme directe de deux sous-espaces:

$$T_xM = X_x \oplus Y_x;$$

2) les champs de plans $X = \{X_x\}$ et $Y = \{Y_x\}$ sont continus e invariants par le difféomorphisme A;

3) pour une métrique riemannienne le difféomorphisme A contracte les plans du premier champ et dilate ceux du second: il existe un $\lambda < 1$ tel que pour tout point $x \in M$ on ait

$$\| A_*\xi \| \leqslant \lambda \| \xi \|, \ \forall \xi \in X_x, \quad \| A_*\eta \| \geqslant \lambda^{-1} \|\eta\|, \ \forall \eta \in Y_x.$$

On dit alors que A est un C-*système*.

E x e m p l e. Soit $M = T^2$ un tore,

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

on automorphisme. Alors A est un C-système.

En effet, les directions propres de l'automorphisme correspondant du plan définissent sur le tore deux champs invariants de directions : l'un contractant, l'autre dilatant.

Remarque 1. On peut remplacer les inégalités précédentes par une condition d'apparence plus faible :

$$\| A_*^n |_X \| \leqslant c\lambda^n, \quad n > 0 ; \quad \| A_*^n |_Y \| \leqslant c\lambda^{-n}, \quad n < 0.$$

Si cette condition est réalisée pour une métrique, elle le sera pour toute autre, éventuellement avec un autre c. Cette condition entraîne les inégalités précédentes (éventuellement pour la métrique modifiée).

2. Cette définition n'implique pas la *différentiabilité* des champs de plans X et Y. Tout difféomorphisme du tore voisin de l'automorphisme

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

est un C-système, cependant ses champs de directions contractant et dilatant peuvent ne pas appartenir à la classe C^2 même s'il est analytique (dans le cas multidimensionnel les champs de plans peuvent ne pas être même de la classe C^1).

3. Le terme C-système et sa définition sont dus à D. Anossov. La lettre C provient du mot *condition*. Anossov a proposé d'appeler les conditions 1 à 3, conditions C et les C-difféomorphismes, C-cascades, Smale a introduit pour eux le terme de « difféomorphismes d'Anossov ».

B. Propriétés des C-difféomorphismes.

Théorème (Anossov). *Tout C-difféomorphisme est structurellement stable.*

La démonstration s'effectue comme au § 13 pour les automorphismes du tore ; on trouvera les détails dans l'article de J. Mather, Anossov diffeomorphisms, Bull. Amer. Math. Soc. 73 (1967), 792-795.

La démonstration liminale utilisait la propriété suivante des C-difféomorphismes.

Théorème. *Les champs de plans contractant et dilatant d'un C-difféomorphisme sont complètement intégrables.*

En d'autres termes, il existe un feuilletage *) contractant et un feuilletage dilatant dont les plans tangents forment les plans contractant et dilatant du difféomorphisme. A noter qu'il est impossible d'utiliser le théorème de Frobénius, puisque les champs ne sont pas différentiables.

---

*) On appelle feuilletage d'une variété de dimension n une décomposition de cette variété en sous-variétés (feuilles) de même dimension k telle que tout point de cette variété possède un voisinage dont la décomposition en composantes connexes des feuilles est difféomorphe à la décomposition d'un cube de dimension n en plans parallèles de dimension k.

La démonstration est basée sur le fait que l'angle des plans pas trop éloignés de ceux du champ dilatant diminue par le C-difféomorphisme : le champ dilatant est un point attractif fixe dans l'espace fonctionnel des champs de plans pour la transformation induite par le C-difféomorphisme dans cet espace.

Pour construire le feuilletage dilatant on peut décomposer la variété en domaines assez petits et considérer dans chacun d'eux un feuilletage quelconque dont les feuilles ont même dimension que les plans du champ dilatant et font avec ces plans un angle pas très élevé. Les images de ces feuilletages par le C-difféomorphisme et ses itérations forment une suite de feuilletages locaux convergeant vers le feuilletage dilatant.

Remarque. Un cas particulier de cette construction est la construction des variétés invariantes rentrante et partante d'un point fixe d un difféomorphisme dont le module de toutes les valeurs propres de la partie linéaire est différent de l'unité. Pour construire la variété partante on peut appliquer les itérations du difféomorphisme à toute variété tangente au sous-espace partant invariant de la partie linéaire du difféomorphisme.

La construction décrite permet d'obtenir les feuilletages contractant et dilatant non seulement du C-difféomorphisme donné mais aussi de tous les C-difféomorphismes voisins. Donc, la propriété d'être C-difféomorphisme n'est pas modifiée par une déformation du difféomorphisme (petite avec les dérivées). D'autre part, la construction montre que les feuilletages (ou mieux les champs de plans dilatant et contractant) dépendent continûment du difféomorphisme.

Une fois que les feuilletages contractant et dilatant pour les difféomorphismes initial et perturbé sont construits, la démonstration du théorème d'Anossov ne présente plus aucune difficulté.

En effet, soit un point de l'espace des phases et la suite de ses images par le difféomorphisme initial. Considérons un système d'ε-voisinages des points images. La distance du difféomorphisme perturbé à l'initial est choisi d'après le nombre petit ε. Si cette distance est assez petite, tout ε-voisinage est décomposé en composantes connexes de feuilles du feuilletage contractant aussi bien pour le difféomorphisme initial que pour le difféomorphisme perturbé.

Nous appellerons ces composantes *disques* verticaux. Considérons le disque vertical du feuilletage passant par l'origine et ses images par les puissances du C-difféomorphisme initial.

Le feuilletage perturbé possède un disque vertical unique dont les images par les puissances positives du difféomorphisme perturbé restent à l'intérieur des ε-voisinages décrits plus haut.

En effet, le C-difféomorphisme initial dilate dans la direction horizontale, donc le difféomorphisme perturbé dilate également dans la direction horizontale.

Appelons U_n les voisinages précédents, $U_n \to B_n$ leurs feuilletages en disques verticaux perturbés, A le difféomorphisme pertur-

bé. Le difféomorphisme A étant dilatant dans la direction horizontale, l'application A^{-1} induit des applications *contractantes* a_n: $B_n \to B_{n-1}$. Le point cherché $b_0 \in B_0$ se définit maintenant comme

$$b_0 = \bigcap_{n \to +\infty} a_1 a_2 \ldots a_n B_n.$$

De façon analogue, il existe un disque perturbé horizontal unique dont les images par les puissances négatives du même difféomorphisme ne quittent pas les voisinages d'indices négatifs.

L'intersection des disques perturbés construits — le disque horizontal et le disque vertical — détermine le point que l'homéomorphisme, qui transforme le C-difféomorphisme non perturbé dans le C-difféomorphisme perturbé, associe au point initial de l'espace des phases.

Désormais il est aisé de vérifier que la construction décrite définit réellement un homéomorphisme conjuguant le C-difféomorphisme non perturbé au difféomorphisme perturbé.

Les C-difféomorphismes à mesure invariante définis par une densité positive possèdent un ensemble partout dense de points périodiques (cycles). D. Anossov et Y. Sinaï (cf. Anossov, Flots géodésiques sur variétés riemanniennes fermées à courbures négatives, Troudi. Matem. in-ta Steklova, **90** (1967), 3-209) ont procédé à une étude assez complète des propriétés ergodiques des C-difféomorphismes à mesure invariante (brassage, etc.).

C. C-flots.

La définition de l'hyperbolicité doit être légèrement revue dans le cas de groupes à un paramètre de difféomorphismes, puisqu'il n'existe ni contraction ni dilatation le long des trajectoires.

Etudions les courbes intégrales de l'équation $\dot{x} = -x$, $\dot{y} = y$ dans le cas d'un col (fig. 83). L'axe t est la trace de deux plans engendrés par les courbes intégrales s' approchant de cet axe pour $t \to +\infty$ (le plan (x, t)) et pour $t \to -\infty$ (le plan (y, t)) ; les autres courbes intégrales s'éloignent de cet axe aussi bien pour $t \to +\infty$ que pour $t \to -\infty$.

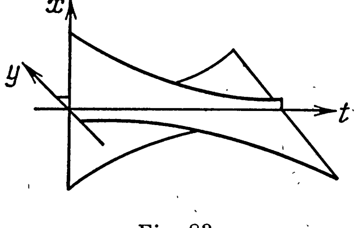


Fig. 83.

Un groupe à un paramètre de difféomorphismes est un C-flot si au voisinage d'une trajectoire quelconque donnée les trajectoires sont disposées comme les courbes intégrales de l'exemple précédent. Voici la définition formelle.

D é f i n i t i o n. Soient M, variété différentiable compacte, v, champ de vecteurs sans singularités sur M, $\{g^t\}$, le flot correspondant. Supposons que

1) tout espace tangent à M est la somme directe des trois sous-espaces

$$T_xM = X_x \oplus Y_x \oplus Z_x;$$

2) les champs de plans X, Y, Z sont continus et invariants par le flot;

3) le plan Z est engendré par le champ de la vitesse;

4) pour c, λ, constantes positives et une métrique riemannienne sur M on a

$$\| g_*^t |_X \| \leqslant ce^{-\lambda t} \text{ pour } t > 0, \quad \| g_*^t |_Y \| \leqslant ce^{\lambda t} \text{ pour } t < 0.$$

Sous ces conditions le flot est dit C-*flot* et l'équation $\dot{x} = v(x)$, C-*système*.

Exemple. Soit une variété M à trois dimensions, que l'on déduit du produit direct d'un tore par l'intervalle $[0, 1]$ en raccordant les tores-faces par un C-automorphisme: on identifie $(x, 1)$ avec $(Ax, 0)$ pour $x \in T^2$,

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Considérons un champ de vecteurs orienté dans le sens du facteur $[0, 1]$ du produit cartésien $T^2 \times [0, 1]$. Ce champ défini initialement dans le produit cartésien se transforme après raccordement en un champ v différentiable (pourquoi ?) sur M.

Le champ v obtenu définit un C-flot sur M.

Théorème. *Tout C-flot est structurellement stable.*

◀ La démonstration s'effectue par la même méthode que pour les C-difféomorphismes (voir travaux cités plus haut). ▶

Tout C-flot à mesure invariante, à densité positive possède une infinité de trajectoires fermées. Donc, même en se restreignant aux champs de vecteurs structurellement stables on ne peut espérer obtenir dans le cas multidimensionnel un tableau aussi simple avec un nombre fini de positions d'équilibre et de cycles que dans le cas de systèmes sur une sphère à deux dimensions.

En 1961 S. Smale a construit les premiers exemples de systèmes structurellement stables ayant une infinité de cycles. Ces exemples ont fait apparaître une divergence exponentielle des trajectoires seulement sur un sous-ensemble fermé de l'espace des phases. De tels ensembles sont dits maintenant hyperboliques. Leur théorie générale a été élaborée plus tard sous l'influence des C-flots.

L'exhibition de ces exemples a bouleversé les conceptions du comportement des trajectoires des systèmes à plusieurs dimensions. Certains se sont empressés d'affirmer que ces résultats n'avaient aucune valeur réelle, puisque même s'ils sont structurellement stables ces systèmes « ne peuvent décrire aucun processus physique réel » du fait de l'instabilité de trajectoires individuelles.

Cependant il existe des cas réels très importants qui apparemment sont le mieux décrits par des systèmes dont les trajectoires sont à divergence exponentielle. Ainsi les phénomènes du type turbulence et les mouvements des particules à choc élastique (par exemple dans les modèles de gaz dont les molécules sont des sphères solides). Un exemple plus simple mais tout aussi réel est le problème du mouvement le long des géodésiques d'une variété de courbure négative. Nous allons traiter le cas le plus simple de ce problème, celui où le mouvement se déroule sur les géodésiques de surfaces à courbure négative constante. Nous aurons besoin de quelques notions de géométrie de Lobatchevsky.

D. Plan de Lobatchevsky.

On appelle *plan de Lobatchevsky* le demi-plan Im $z > 0$ muni de la métrique *)

$$ds^2 = \frac{dx^2 + dy^2}{y^2}, \quad \text{où} \quad z = x + iy.$$

La droite $y = 0$ est dite *droite absolue*. A noter que dans cette métrique les angles coïncident avec les angles euclidiens et la distance à l'absolu est infinie.

Théorème. *Les cercles et droites orthogonaux à l'absolu et eux seuls sont les géodésiques du plan de Lobatchevsky* (fig. 84).

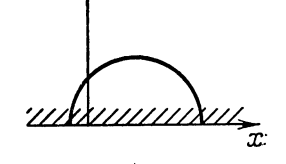


Fig. 84.

◀ La métrique est invariante par

1) les translations le long de l'absolu;

2) les homothéties de centre l'origine des coordonnées;

3) la symétrie $z \mapsto -\bar{z}$ (c'est évident).

On vérifie immédiatement que

4) la métrique est invariante par l'inversion $z \mapsto 1/\bar{z}$.

1) à 4) entraînent l'invariance de la métrique par les applications homographiques du demi-plan supérieur dans lui-même. De plus, de 3) il suit que l'axe y est une géodésique. Or l'image de l'axe y par une application homographique réelle est un cercle ou une droite orthogonaux à l'absolu. Donc tous les cercles et droites orthogonaux à l'absolu sont des géodésiques.

Réciproquement par tout point et dans toute direction il passe un cercle ou une droite orthogonale à l'absolu. Donc il n'existe pas d'autres géodésiques. ▶

Remarque. On a prouvé par la même occasion que les mouvements (qui conservent la métrique et l'orientation) du plan de Lobatchevsky sont des applications homographiques du demi-plan supérieur dans lui-même.

*) Ou toute variété riemannienne isométrique à ce demi-plan.

T h é o r è m e. *Les cercles du plan de Lobatchevsky sont les cercles euclidiens ne coupant pas l'absolu et eux seuls.*

◀ Soit le cercle unité. Par une application homographique on peut envoyer le demi-plan supérieur dans le cercle unité (comparer avec le chapitre premier, § 5, point E). Donc le disque $|z| < 1$ peut être considéré comme un modèle de plan de Lobatchevsky (cf. fig. 32).

Les images des applications homographiques préservant le demi-plan supérieur sont des applications homographiques préservant le cercle unité. Donc, la métrique du plan de Lobatchevsky sur le disque unité est invariante par les applications homographiques préservant le disque.

Or parmi ces applications il existe des rotations de centre O. Donc tous les points d'un cercle euclidien possédant le même centre que le cercle unité sont équidistants de ce centre au sens de la métrique de Lobatchevsky. Ainsi, tous les cercles euclidiens, de même centre que le cercle unité, sont cercles de Lobatchevsky.

Or tout cercle euclidien ne coupant pas la droite absolue est susceptible par un déplacement du plan de Lobatchevsky d'être transformé en cercle euclidien centré en l'origine des coordonnées. Donc, tout cercle euclidien ne coupant pas l'absolu est un cercle au sens de la métrique de Lobatchevsky (aussi bien dans le disque unité que dans le demi-plan). D'où la réciproque: tout cercle de Lobatchevsky est cercle euclidien. ▶

D é f i n i t i o n. On appelle *horicycle* la limite d'une suite de cercles tangents en un point du plan de Lobatchevsky.

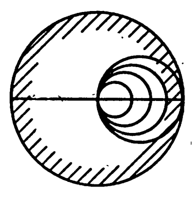


Fig. 85

T h é o r è m e. *Les droites et les cercles euclidiens tangents à l'absolu, et eux seuls, sont horicycles du plan de Lobatchevsky.*

◀ Soit une demi-géodésique issue d'un point du plan de Lobatchevsky (fig. 85). Repérons sur cette demi-géodésique le point qui se trouve à une distance t du point initial. Le cercle de rayon t centré au point repéré est orthogonal à la géodésique au point initial. Supposons que $t \to +\infty$. Au sens euclidien le cercle tracé tend vers le cercle orthogonal à la géodésique envisagée et passant par le point de cette dernière sur l'absolu. Ce cercle euclidien est tangent à l'absolu. ▶

R e m a r q u e 1. Par un passage à la limite analogue pour $t \to \infty$ on construit les horicycles sur des surfaces de courbure négative et les horisphères sur des variétés de courbure négative.

R e m a r q u e 2. Par chaque point du plan de Lobatchevsky il passe deux horicycles ayant une tangente commune; on les obtient à l'aide de la construction précédente pour $t \to +\infty$ et pour $t \to -\infty$.

E. Flots géodésiques sur des surfaces de courbure négative.

Soit M une variété riemannienne. On admettra que M est complète comme un espace métrique. Par exemple, toute variété compacte est complète; le plan de Lobatchevsky est complet car la distance à l'absolu est infinie.

Considérons l'ensemble de tous les vecteurs de longueur 1 tangents à la variété M. Cet ensemble est une variété de dimension $2n - 1$, n étant la dimension de M. On le désigne par T_1M.

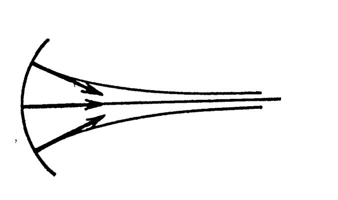


Fig. 86

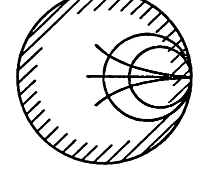


Fig. 87

D é f i n i t i o n. On appelle *flot géodésique* sur M le groupe à un paramètre de difféomorphismes de la variété T_1M, défini comme suit: tout vecteur avance d'une distance t pendant le temps t le long de la géodésique à laquelle il est tangent, en lui restant tangent.

T h é o r è m e. *Le flot géodésique sur le plan de Lobatchevsky vérifie les conditions* 1) *à* 4) *de la définition d'un* C-*flot.*

◀ 1°. Construisons les feuilletages contractant et dilatant. A cet effet, associons à tout vecteur l'horicycle qui lui est orthogonal et qui est la limite de tous les cercles dont les centres sont situés en avant du point d'application de ce vecteur. Traçons un vecteur unitaire normal à l'horicycle en chacun de ses points. On obtient ainsi un champ continu de vecteurs normaux (fig. 86).

A signaler que si nous avions commencé par l'un de ces vecteurs, nous aurions débouché sur le même horicycle avec le même champ. Cet horicycle à champ peut être regardé comme une courbe dans l'espace à trois dimensions T_1M des vecteurs unitaires tangents du plan de Lobatchevsky. Nous avons donc construit des feuilletages de dimension un dans T_1M, c'est-à-dire nous avons partitionné l'espace des vecteurs unitaires tangents en courbes. Cette partition est le feuilletage contractant.

Le feuilletage dilatant se construit de façon analogue par la considération des cercles dont les centres sont situés en arrière du point d'application du vecteur.

2°. Les conditions 2 et 3 qui sont immédiates à vérifier traduisent l'invariance des géodésiques et des horicycles par le flot géodésique.

En effet, toute famille de géodésiques orthogonales à un horicycle coupe l'absolu au point de contact de l'absolu et de cet horicycle; tout horicycle tangent à l'absolu en ce point de contact est orthogonal à toutes les géodésiques de la famille (fig. 87).

Donc le flot géodésique associe à chaque horicycle (muni d'un champ de normales) l'horicycle tangent à l'absolu au même point (et muni également d'un champ de normales).

3°. La condition 1 revient à dire que les vecteurs tangents à une géodésique munie d'un champ tangent et aux deux horicycles munies de champs normaux sont linéairement indépendants. La vérification est immédiate: l'essentiel est que le contact des deux horicycles soit du premier ordre et non pas d'ordre supérieur.

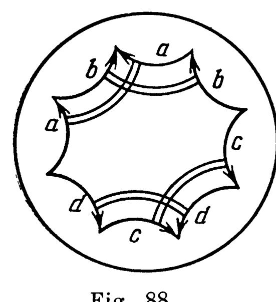


Fig. 88

4°. Montrons que le flot contracte exponentiellement des portions de l'horicycle contractant. Supposons que l'horicycle initial est la droite $y = 1$ dans le demi-plan $y > 0$. Les géodésiques sont les droites $x = \text{const}$; le flot géodésique transforme la droite $y = 1$ en la droite $y = e^t$ pendant le temps t (la distance du point 1 au point y est égale à $\ln y$ le long de l'axe y). Par suite, tout intervalle d'horicycle se transforme en un intervalle e^t fois plus petit. D'où il suit que le flot contracte les feuilles du feuilletage contractant (au sens de la métrique naturelle de T_1M).

Par un raisonnement analogue on achève de vérifier la condition 4 de la définition d'un C-système pour les horicycles dilatants. ▶

C o r o l l a i r e. *Un flot géodésique sur une surface compacte de courbure négative constante est un* C-*flot.*

◀ Par un changement horaire on ramène cette surface à une surface de courbure —1. Le revêtement universel d'une telle surface est le plan de Lobatchevsky. Elle se déduit à partir de ce dernier par identification des points homologues par un groupe discret de mouvements du plan de Lobatchevsky (fig. 88).

Par cette identification les géodésiques, cercles, horicycles du plan de Lobatchevsky se projettent respectivement sur les géodésiques, les cercles, les horicycles de la surface; le flot géodésique sur la surface et ses feuilletages contractant et dilatant se projettent sur des feuilletages de même nature pour la surface. ▶

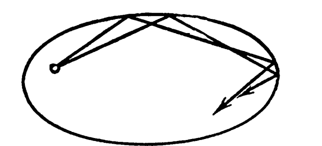
De ce qui précède il suit, en particulier, qu'un *flot géodésique sur une surface compacte de courbure négative constante est structurellement stable et possède un ensemble partout dense de géodésiques fermées.*

R e m a r q u e. Un flot géodésique sur une variété multidimensionnelle de courbure négative (pas nécessairement constante) est également C-flot. La démonstration diffère peu de celle du cas élé-

mentaire: seul l'établissement de l'existence des horicycles (horisphères) est un peu plus compliqué. Voir l'ouvrage cité plus haut de D. Anossov.

F. Systèmes à chocs élastiques.

Soit un flot géodésique sur la surface d'un ellipsoïde. Si le petit axe de l'ellipsoïde tend vers zéro, celui-ci s'aplatit et tend vers une ellipse. Le flot géodésique se transforme à la limite en un système dit *système à chocs élastiques* dans le domaine borné par l'ellipse: le



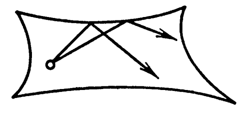


Fig. 89

point courant se déplace sur une droite à l'intérieur du domaine et se réfléchit sur la frontière selon la loi « l'angle d'incidence est égale à l'angle de réflexion » (fig. 89).

Une trajectoire à chocs élastiques n'est jamais partout dense à l'intérieur d'une ellipse. Cependant à l'intérieur de domaines bornés par d'autres courbes (par exemple des courbes non différentiables de convexité tournée vers l'intérieur) le mouvement à chocs élastiques possède presque les mêmes propriétés d'instabilité exponentielle des trajectoires et de brassage que les C-flots.

Soit, en particulier, un système à chocs élastiques sur un tore troué. Ce système peut être considéré comme la limite des flots géodésiques sur la surface de genre deux (qui dégénère en un tore bilatéral troué au même titre qu'un ellipsoïde en une ellipse bilatérale). Bien plus, on peut considérer que le tore bilatéral troué muni d'une métrique plane est la limite d'une surface de genre deux de courbure négative (toute la courbure se réunit par dégénérescence sur le bord du trou). Aussi n'est-il pas étonnant que ce système soit doué des propriétés d'un C-flot.

Il reste à espérer que des raisonnements proches de la théorie hyperbolique permettent de prouver l'ergodicité d'un système de boules solides dans une boîte, postulé par Boltzmann en mécanique statistique. (L'ergodicité signifie que tout sous-ensemble invariant de l'espace des phases est de mesure nulle, ou de mesure complète; l'ergodicité entraîne la coïncidence presque partout des moyennes temporelles et spatiales. L'espace des phases est un ensemble de niveaux d'énergie.) Dans le cas plan la démonstration a été publiée par Y. Sinaï (Y. Sinaï, Systèmes dynamiques à réflexions élastiques.

Propriétés ergodiques des billards dispersants, UMN 25, 2 (1970), 141-192). On pourra consulter aussi L. Bounimovitch, Sur les billards proches des dispersants, Matem. sbornik 94, 1 (1974), 49-73.

G. C-systèmes et balayage.

Le cas hyperbolique se présente dans les problèmes d'analyse numérique résolubles par la méthode de balayage. Supposons par exemple que nous ayons à résoudre le problème aux limites pour l'équation du second ordre $\ddot{x} = x$ (c'est-à-dire pour le système $\dot{x} = p$, $\dot{p} = x$) sur l'intervalle $[0, T]$. Supposons données les conditions aux

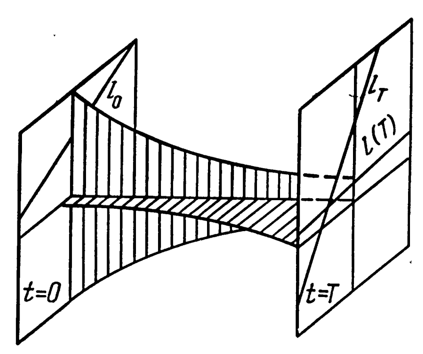


Fig. 90

limites non homogènes: le point initial $\varphi(0)$ de coordonnées $(x(0), p(0))$ est situé sur une droite donnée l_0 du plan (x, p) et l'extrémité $\varphi(T)$ sur une droite donnée l_T.

Si le point initial $\varphi(0)$ était connu, la résolution du problème de Cauchy avec la condition initiale $\varphi(0)$ donnerait lieu à une perte de précision croissant suivant la loi exponentielle avec la longueur T de l'intervalle d'intégration. En effet, les solutions vérifiant une condition initiale proportionnelle au vecteur dilatant $(1, 1)$ sont exponentiellement croissantes. Donc, le passage du plan $t = 0$ au plan $t = T$ donne lieu à une dilatation dans le sens du vecteur $(1, 1)$ (dans la suite, la direction de ce vecteur sera appelée horizontale) et une contraction dans le sens du vecteur $(1, -1)$ (resp. la direction verticale, cf. fig. 90).

Etudions maintenant l'image de la droite l_0 par l'application considérée. Si l'image de tout point de la droite est déterminée avec une perte de précision exponentiellement croissante, celle de la droite par contre l'est exactement.

En effet, la direction de cette image est généralement proche de l'horizontale.

Par conséquent, une erreur dans le calcul d'un point de cette droite pratiquement horizontale se répercute peu sur la position

de cette droite: l'erreur peut posséder une composante horizontale élevée et une composante verticale petite.

Le point $\varphi(T)$ est l'intersection de la droite l_T et de l'image de la droite l_0. Il ne reste plus maintenant qu'à résoudre le problème de Cauchy regressivement. On remarquera que l'erreur sur la composante horizontale ne croît plus, tandis que la composante verticale du point $\varphi(t)$ est fixe du fait qu'elle est située sur la droite $l(t) = g^t l_0$. Donc, en allant de 0 à T on détermine les droites $l(t)$, puis en revenant de T à 0 on repère un point sur chacune de ces droites et ce sans perte exponentielle de précision.

H. Domaines d'applications des C-systèmes.

Il est des C-systèmes et des objets similaires comme des cycles limites du temps de Poincaré. L'appareil mathématique des cycles limites avait été entièrement mis au point mais on ne commença à s'y intéresser sérieusement que plusieurs décennies plus tard lorsque le développement de la radiotechnique fit de la théorie des oscillations non linéaires un domaine des mathématiques appliquées.

L'hypothèse que la théorie du mouvement turbulent d'un liquide est le champ naturel d'application des C-flots est connue depuis le début des années 1960. Imaginons-nous un récipient clos, rempli d'un liquide incompressible visqueux, mis en mouvement par une force extérieure (un agitateur). Le rôle de l'agitateur est de s'opposer à l'effet de la viscosité et de maintenir le liquide en mouvement.

Les équations hydrodynamiques de Navier-Stokes définissent un système dynamique *) dans un espace fonctionnel (les points de cet espace des phases de dimension infinie sont les champs de vecteurs de divergence zéro, les champs de vitesses du liquide).

Les positions d'équilibre de ce système dynamique sont les champs stationnaires de vitesses, c'est-à-dire des mouvements de liquide dont la vitesse est invariable avec le temps. Les cycles de ce système correspondent à des mouvements périodiques du liquide dont la vitesse varie périodiquement en chaque point de l'espace. On observe parfois un tel mouvement en ouvrant un robinet d'eau.

L'hypothèse de la description mathématique de la turbulence consiste en la réductibilité du problème à un système dynamique de dimension finie puisque la viscosité amortit rapidement les harmoniques élevés. En d'autres termes, on suppose que l'espace des phases de dimension infinie contient une variété ou un ensemble de dimension finie vers lequel tendent toutes les trajectoires ; sur cet ensemble le flot est un C-système ou est doué de propriétés proches de l'instabilité exponentielle des trajectoires et du brassage.

---

*) En toute rigueur la théorie des équations aux dérivées partielles n'a toujours pas réglé le problème de l'existence et de l'unicité des solutions de l'équation de Navier-Stokes à trois dimensions, mais nous négligerons ce fait.

Dans ce cas le mouvement du liquide doit posséder les propriétés suivantes: pour toute condition initiale le mouvement doit rapidement prendre un régime défini ; cependant ce régime n'est ni stationnaire ni périodique; même si le régime limite est déterminé par un nombre fini de paramètres (les « phases» de ce régime), ces paramètres sont extrêmement instables (les courants limites de phases initiales voisines divergent) ; du reste les caractéristiques statistiques du mouvement ne dépendent pas de ces phases instables.

Voici ce qui a été fait dans cette direction. Si la viscosité est assez élevée, le système de Navier-Stokes possède un point fixe unique attractif pour toutes les trajectoires. Ce mouvement est dit laminaire. Tout autre mouvement a tendance à se transformer en mouvement laminaire sous l'influence de la viscosité. Avec une diminution de la viscosité le mouvement laminaire risque de perdre sa stabilité, ce qui donnerait naissance à un cycle limite stable (cf. chapitre 6). Si la viscosité baisse encore, le cycle est susceptible de perdre sa stabilité et de se transformer en un mouvement plus complexe, non périodique attirant les mouvements voisins. Tout laisse à croire que les trajectoires de ce mouvement sont généralement exponentiellement instables sur un ensemble attractif. Cette hypothèse n'a encore rien d'un théorème même si ces dernières années elle fait l'objet de nombreuses recherches aussi bien théoriques qu'expérimentales (cf. par exemple J. B. Mclaughlin, P. C. Martin, Transition to turbulence of a statically stressed fluid system, Phys. Rev. A **12** (1975), 186-203).

A signaler du reste que l'appartion d'un ensemble attractif contenant des trajectoires exponentiellement instables n'est pas forcément liée à la perte de stabilité du mouvement laminaire: cet ensemble peut très bien apparaître loin de la position d'équilibre et voire même pour des valeurs de la viscosité telles que le mouvement laminaire est encore stable.

§ 15. Systèmes structurellement stables non partout denses

Dans ce paragraphe on exhibe un domaine de l'espace fonctionnel des systèmes dynamiques de la classe C^1 ne contenant pas de systèmes structurellement stables.

A. Exemple de Smale.

En 1965, S. Smale construisit un exemple de difféomorphisme du tore à trois dimensions, dont le voisinage ne renfermait aucun difféomorphisme structurellement stable.

Par conséquent, sur une variété à quatre dimensions il existe un champ de vecteurs qui ne peut être rendu structurellement stable par une petite déformation.

Par la suite de tels champs ont été construits sur des variétés à trois dimensions (cf. S. Newhous, Nondensity of Axiom A(a), Global Analysis, Proceed. Simp. Pure Math. AMS, **14** (1971), 191-203). Dans ce paragraphe on expose la constructon de Smale.

B. Description de l'exemple.

Munissons T^3 des coordonnées (x, y, z mod 2π). Nous allons définir un difféomorphisme A: $T^3 \to T^3$ aux voisinages du tore T^2: $z = 0$ et d'un intervalle de l'axe z (la forme de A importe peu dans le reste du tore).

Au voisinage U du tore T^2 l'application A est donnée par

$$A\ (x,\ y;\ z) = (2x + y,\ x + y;\ z/2).$$

Au voisinage du point $O\,(0,\ 0,\ \pi)$ l'application A a pour formule

$$A\ (x,\ y;\ \pi + u) = (x/2,\ y/2;\ \pi + 2u).$$

Donc le point O est un col, et la variété invariante partante est une courbe γ contenant l'intervalle $(\pi,\ \pi - \varepsilon)$ de l'axe z.

La courbe γ est invariante et dilatée par A ; donc les puissances de A envoient l'intervalle envisagé dans la moitié d'une variété invariante qui, soit a pour extrémité le point fixe de l'application A, soit est de longueur infinie.

On exigera que la courbe pénètre dans le domaine U et y soit de longueur illimitée. Il est aisé de voir que les difféomorphismes du tore doués des propriétés indiquées existent.

C. Propriétés stables du difféomorphisme A.

1°. *La restriction de A à un voisinage assez petit du tore T^2 est structurellement stable.*

◀ On démontre cette proposition en utilisant le procédé qui a servi à prouver le théorème de Grobman-Hartman. Remplaçons le difféomorphisme $A: T^3 \to T^3$ par l'application $A': T^2 \times \mathbb{R} \to T^2 \times \times\, \mathbb{R}$, définie partout par la même formule que celle de $\tilde{A}$ dans U. Le difféomorphisme A proche de $\tilde{A}$ peut être remplacé par l'application $\tilde{A}': T^2 \times \mathbb{R} \to T^2 \times \mathbb{R}$ adaptée à $\tilde{A}$ dans un voisinage du tore $T^2 \times O$, de sorte que la différence $\tilde{A}' - A'$ possède un support compact et est petite dans C^1. Désormais nous pouvons nous servir du théorème d'Anossov (ou, plus exactement, de sa démonstration) ; on établit l'équivalence topologique de $\tilde{A}'$ et A' et, partant, celle de $\tilde{A}$ et A dans un voisinage du tore T^2. ▶

2°. De ce qui précède il suit que le difféomorphisme A possède une variété invariante $\tilde{T}^2$ proche du tore T^2, et sur elle un ensemble dénombrable partout dense de points périodiques. Par chaque point du voisinage $\tilde{U}$ du tore $\tilde{T}^2$ passe une feuille différentiable définie de façon unique et dépendant continûment d'un point du feuilletage

contractant à deux dimensions du difféomorphisme $\tilde{A}$ (ce feuilletage est composé de points qui se rapprochent par les puissances de $\tilde{A}$).

3°. *L'application $\tilde{A}$ possède un point fixe $\tilde{O}$ voisin du point fixe O de l'application A.*

◀ Ceci est une conséquence du théorème des fonctions implicites, puisque le point $\tilde{O}$ est non dégénéré et le difféomorphisme $\tilde{A}$ proche de A. ▶

Les valeurs propres du linéarisé de $\tilde{A}$ au point $\tilde{O}$ sont voisines des valeurs propres de A en O.

D'après le théorème de Grobman-Hartman, le point $\tilde{O}$ (comme le point O) est un col d'où est issue une variété invariante partante $\tilde{\gamma}$ proche, ce qu'on voit aisément, de γ.
En particulier, $\tilde{\gamma}$ entre dans le voisinage $\tilde{U}$ du tore $\tilde{T}^2$.

4°. Par une déformation aussi petite que l'on veut de l'application A loin de U on peut modifier la courbe de telle sorte qu'elle présente un « nez »: localement elle sera située d'un côté d'une feuille du feuilletage contractant de A contenant un point de la courbe γ, le contact avec γ étant du premier ordre. Désignons par A_1 l'application ainsi définie.

D. Instabilité structurelle.

Prouvons que *le difféomorphisme A_1 est structurellement instable avec tous les difféomorphismes voisins.*

◀ Incluons l'application A_1 dans une famille à un paramètre de difféomorphismes A_s différant peu de A_1 au voisinage de l'antécédent d'un point du nez. Chaque application de la famille A_s possède les propriétés de A_1: tore invariant, feuilletage contractant à deux dimensions, col à variété invariante partante présentant un nez. Nous admettrons que lorsque s varie le nez se déplace transversalement aux feuilles du feuilletage contractant.

Considérons maintenant la feuille contractante contenant le nez. Cette feuille est susceptible de contenir ou de ne pas contenir un point périodique du tore. Les points périodiques étant partout denses dans le tore, par une déformation aussi petite que l'on veut de A_1 on peut transporter le nez sur une feuille contenant ou ne contenant pas un point périodique.

Or la propriété d'une feuille de contenir ou de ne pas contenir un point périodique est topologiquement invariante, donc le type topologique de A_1 varie par une déformation aussi petite que l'on veut de A_1. Par conséquent, l'application A_1 est structurellement instable.

Soit $\tilde{A}_1$ un difféomorphisme proche de A_1. D'après ce qui a été dit au point C, on peut effectuer pour $\tilde{A}_1$ la même construction que pour A_1. Donc, $\tilde{A}_1$ est un difféomorphisme structurellement instable. ▶

CHAPITRE 4

THÉORIE DES PERTURBATIONS

La plupart des équations différentielles non seulement ne peuvent pas être résolues analytiquement, mais encore ne donnent lieu à aucune étude qualitative tant soit peu complète. La théorie des perturbations est un arsenal au plus haut point utile de méthodes de résolution des équations proches d'équations spéciales — dites équations non perturbées — dont les solutions sont connues. La théorie des perturbations tient compte de l'influence de faibles modifications des équations différentielles sur le comportement des solutions.

Si la perturbation est de l'ordre de ε, son influence sur des durées de temps de l'ordre de 1 modifie la solution d'une quantité de l'ordre de ε. On détermine cette quantité approximativement par résolution des équations aux variations sur la solution non perturbée. Mais si l'on s'intéresse au comportement de la solution sur un long intervalle de temps, de l'ordre de $1/\varepsilon$ par exemple, on est confronté à un problème bien plus complexe qui constitue l'objet des méthodes asymptotiques d'étude de la théorie des perturbations. La plus importante de ces méthodes est la méthode de moyennisation que l'on se propose de développer dans ce chapitre.

La méthode de moyennisation est utilisée en mécanique céleste depuis les temps de Lagrange et Laplace pour déterminer l'évolution des orbites planétaires sous l'influence des perturbations mutuelles des planètes. Gauss en a donné la définition suivante: pour déterminer l'évolution il faut distribuer la masse de chaque planète sur son orbite proportionnellement au temps passé dans chaque partie de l'orbite et remplacer l'attraction des planètes par celle des anneaux obtenus.

Cependant la justification de la méthode de moyennisation est un problème qui est encore loin d'être résolu entièrement.

§ 16. Méthode de moyennisation

Dans ce paragraphe on décrit la variante la plus simple de la méthode de moyennisation. Sa justification fera l'objet des paragraphes suivants.

A. Systèmes non perturbé et perturbé.

Soit un fibré différentiable $\pi: M \to B$. Un champ de vecteurs v sur la variété M est *vertical* s'il est tangent à chaque fibre (fig. 91). Les fibres sont généralement des tores dans les applications.

Les fonctions sur la base B du fibré π sont intégrales premières de l'équation $\dot{x} = v(x)$ sur M. Le champ de vecteurs vertical v est dit *non perturbé*. Le champ $v + \varepsilon v_1$ proche de v est dit *champ perturbé*. Soit l'équation différentielle perturbée

$$\dot{x} = v(x) + \varepsilon v_1(x).$$

Toute trajectoire de l'équation non perturbée se projette par l'application π en un point de la base B. Tout mouvement sur une trajectoire de l'équation perturbée est associé par π à un mouvement

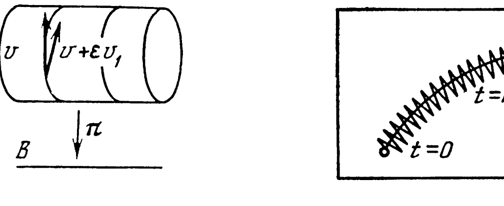


Fig. 91 Fig. 92

dont la vitesse est de l'ordre de ε. On constate un sensible déplacement de la projection sur la base pendant une durée de temps de l'ordre de $1/\varepsilon$. La méthode de moyennisation est destinée à étudier ce mouvement lent sur la base à l'aide d'un champ de vecteurs. Ce mouvement lent est décrit comme la combinaison de petites oscillations avec évolution systématique ou dérive (fig. 92).

E x e m p l e. Considérons le système planétaire. Les équations non perturbées tiennent compte seulement de l'interaction du soleil et des planètes. Dans les mouvements non perturbés, les planètes se déplacent sur des orbites képlériennes. Ici la perturbation est l'attraction mutuelle des planètes. ε est égal au rapport de la masse des planètes à celle du Soleil; c'est une quantité de l'ordre de $1/1000$. [L'unité caractéristique de temps est la période de révolution autour du Soleil c'est-à-dire une quantité de l'ordre d'une année ou de dix ans.] L'unité de temps est la période de révolution autour du Soleil, c'est-à-dire une unité de l'ordre d'une année ou d'une décennie, l'unité de longueur est le rayon de l'orbite de la planète.

Dans cet exemple M est l'espace des phases, la base B, l'espace des ellipses képlériennes, les fibres, des tores de dimension égale au nombre de planètes (chaque collection d'ellipses képlériennes définit un tore dont tout point est déterminé par la donnée des positions des planètes sur les ellipses). Un déplacement de l'ordre de 1 sur la

base correspond ainsi à une réduction de moitié du rayon de l'orbite par exemple. La quantité $1/\varepsilon$ représente une durée de l'ordre du millénaire ou de dix millénaires.

Donc, la dérive sur la base à une vitesse de ε peut dans l'exemple considéré réduire de moitié le rayon de l'orbite de la Terre, ce qui aurait des conséquences catastrophiques pour notre civilisation qui doit son existence précisément à l'absence de dérive (du moins en ce qui concerne les rayons des orbites; l'excentricité des orbites varie et cette variation se répercute probablement sur les périodes glaciaires).

B. Procédure de moyennisation.

Pour décrire la méthode de moyennisation nous aurons besoin des notations suivantes. On supposera que les fibres du fibré sont des tores à n dimensions. Le fibré est un produit direct au voisinage de tout point de la base. On retiendra un tel voisinage et on se donnera tout point de l'espace du fibré M par un couple (φ, I), où I est un point de la base, φ un point d'un tore F à n dimensions.

Le choix de la notation I est dû au fait que les coordonnées $(I_1, \ldots, I_k)$ du point I définissent sur M les intégrales premières du système non perturbé. Un point φ du tore F est défini par n coordonnées angulaires $(\varphi_1, \ldots, \varphi_n)$ mod 2π. [Dans les applications les coordonnées φ_k sont généralement définies de façon unique à l'origine près sur chaque tore et aux applications linéaires unimodulaires entières près.] Nous fixons le système de coordonnées (φ, I).

D é f i n i t i o n. On appelle *équation non perturbée* de la méthode de moyennisation l'équation

$$\begin{cases} \dot{\varphi} = \omega(I), \\ \dot{I} = 0, \end{cases}$$

où ω est un champ de vecteurs vertical défini par le vecteur de fréquences $(\omega_1(I), \ldots, \omega_n(I))$ qui dépend d'un point I de la base.

D é f i n i t i o n. On appelle *équation perturbée* de la méthode de moyennisation l'équation

$$\begin{cases} \dot{\varphi} = \omega(I) + \varepsilon f(I, \varphi, \varepsilon), \\ \dot{I} = \qquad\quad \varepsilon g(I, \varphi, \varepsilon), \end{cases}$$

où f et g sont 2π-périodiques sur φ, $\varepsilon \ll 1$ un petit paramètre. Les coordonnées angulaires φ_i sont dites *variables rapides*, les coordonnées I_j, *variables lentes.*

D é f i n i t i o n. On appelle *équation moyennisée* l'équation

$$\dot{J} = \varepsilon G(J),$$

où $G(J) = \oint g(J, \varphi, 0)\, d\varphi \Big/ \oint d\varphi$ est la valeur moyenne de la fonction g sur une fibre.

Les solutions de l'équation moyennisée sont dites *mouvements moyens.*

E x e m p l e. Soit l'équation perturbée

$$\dot{\varphi} = \omega, \quad \dot{I} = \varepsilon (a + b \cos \varphi).$$

L'équation moyennisée s'écrit

$$\dot{J} = \varepsilon a.$$

Donc, quand on passe à l'équation moyennisée on élimine une quantité du même ordre que celles qui restent au second membre de l'équation en I. Sur des durées de temps de l'ordre de 1 les quantités éliminées fournissent le même effet (de l'ordre de ε) que les quantités

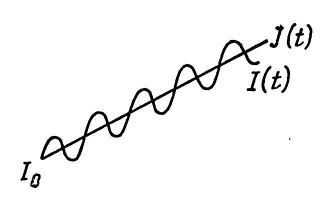


Fig. 93

restantes. Mais sur des durées de l'ordre de $1/\varepsilon$ cet effet est foncièrement différent: les termes restants entraînent une dérive systématique, les termes éliminés, une petite pulsation.

◀ La solution de l'équation perturbée s'écrit (pour $\varphi_0 = 0$ par exemple)

$$I(t) = I_0 + \varepsilon\, at + \varepsilon b \sin \omega t/\omega.$$

On remarque que cette solution ne diffère de celle de l'équation moyennisée, $J(t) = I_0 + \varepsilon at$ (fig. 93) que d'un petit terme oscillant. ▶

C. Moyennes spatiale et temporelle.

Soit un intervalle de temps T, grand en regard de 1, et petit en regard de $1/\varepsilon$. Au cours de cet intervalle la trajectoire du mouvement perturbée s'éloigne peu de la fibre initiale.

Calculons le chemin parcouru par la projection de la trajectoire perturbée sur la base pendant le temps T. Ce chemin est une quantité de l'ordre de $\varepsilon T \ll 1$. La vitesse est égale à $\varepsilon g(I, \varphi, \varepsilon)$. On peut admettre en première approximation que I est constant, ε nul et que les variations de φ obéissent à l'équation non perturbée. L'expression

approchée du chemin parcouru pendant le temps T est alors

$$\Delta I = \varepsilon T \left[\frac{1}{T} \int_0^T g(I, \varphi(t), 0)\, dt \right] + o(\varepsilon T).$$

Le temps $T \gg 1$ est élevé, donc la quantité entre crochets est proche de la moyenne temporelle de la fonction g.

Soit le temps lent $\tau = \varepsilon t$. Lorsque t varie de 0 à $1/\varepsilon$, τ varie de 0 à 1. On désignera la vitesse des mouvements par rapport au temps moyen par le symbole prime. L'égalité précédente s'écrit alors

$\frac{\Delta I}{\Delta\tau} \approx$ moyenne temporelle de g, $I' =$ moyenne temporelle de g.

Remplaçons la moyenne temporelle par la moyenne spatiale. On obtient l'équation moyennisée.

$$J' = G(J), \quad G = \text{moyenne spatiale de } g.$$

Donc, le passage à l'équation moyennisée revient à remplacer les moyennes temporelles par les moyennes spatiales le long du mouvement non perturbé.

D. Discussion.

La méthode de moyennisation consiste à remplacer l'équation perturbée par une équation moyennisée bien plus simple, dont les solutions sont étudiées sur des intervalles de temps de l'ordre de $1/\varepsilon$ (c'est-à-dire sur des intervalles de temps lent de l'ordre de 1). Ensuite on tire des conclusions sur le comportement du mouvement perturbé pendant des intervalles de temps de l'ordre de $1/\varepsilon$ (généralement ces conclusions consistent à établir que la I-composante de la solution de l'équation perturbée est voisine de la solution de l'équation moyennisée pendant la durée $1/\varepsilon$).

Cette conclusion n'est pas une conséquence des raisonnements précédents et doit être justifiée. En effet, pour déduire l'équation moyennisée on a remplacé les moyennes temporelles par les moyennes spatiales. Cette substitution ne présente de l'intérêt que si la trajectoire du mouvement non perturbé est uniformément distribuée sur le tore à n dimensions, c'est-à-dire lorsque les fréquences sont incommensurables. Cependant dans le cas résonnant la trajectoire du mouvement non perturbé est partout dense non pas dans un tore à n dimensions, mais dans un tore de dimension inférieure. Donc la substitution de la moyenne spatiale à la moyenne temporelle sur un tore à n dimensions au voisinage des résonances est manifestement illicite. En effet, il existe des exemples qui montrent que la différence entre la projection de la trajectoire perturbée sur la base et la solution de l'équation moyennisée atteint une valeur de l'ordre de 1 pendant le temps $1/\varepsilon$: la dérive moyenne et la projection du mouvement réel ne sont pas orientées dans le même sens.

L'unique cas à être étudié jusqu'au bout est celui des systèmes à une fréquence où les fibres sont des tores à une dimension, c'est-à-dire des cercles.

§ 17. Moyennisation dans les systèmes à une fréquence

On énonce et on démontre le théorème justifiant la méthode de moyennisation pour les systèmes à une fréquence.

A. Enoncé du théorème.

Soit un espace des phases M, produit direct d'un domaine B de l'espace euclidien $\mathbb{R}^k$ et d'un cercle S^1. Désignons par $\varphi \bmod 2\pi$ la coordonnée angulaire sur le cercle, par I un point de la base B.

L'équation perturbée

$$\dot{\varphi} = \omega(I) + \varepsilon f(I, \varphi, \varepsilon),$$
$$\dot{I} = \varepsilon g(I, \varphi, \varepsilon),$$

où f et g sont des fonctions 2π-périodiques en φ, nous donne l'équation moyennisée

$$\dot{J} = \varepsilon G(J), \quad G(J) = \frac{1}{2\pi}\int_0^{2\pi} g(J, \varphi, 0)\, d\varphi.$$

Considérons un point initial I_0 de B et supposons que la solution $J(t)$ de l'équation moyennisée vérifiant la condition initiale $J(0) = I_0$ reste dans le domaine B pendant le temps $t = T/\varepsilon$ (c'est-à-dire que la solution de l'équation $dJ/d\tau = G(J)$ vérifiant la condition initiale I_0 ne quitte pas le domaine B pendant le temps lent $\tau = T$).

Théorème. *Supposons que la fréquence ω ne s'annule pas dans le domaine B. Alors l'écart entre la valeur de la solution de l'équation moyennisée $J(t)$ et la I-composante de la solution $I(t)$ de l'équation perturbée vérifiant la condition initiale $I(0) = J(0)$ reste petit pendant le temps $t \in [0, T/\varepsilon]$, pourvu que ε soit suffisamment petit :*

$$|I(t) - J(t)| \leqslant C\varepsilon,$$

où la constante C ne dépend pas de ε.

B. Construction principale.

L'idée maîtresse de la démonstration du théorème est d'éliminer la perturbation par un changement de variables convenable. Cette idée connaît plusieurs développements (voir, par exemple, le paragraphe précédent et le suivant) et constitue la voûte de l'appareil formel tout entier de la théorie des perturbations.

Remplaçons I par une nouvelle coordonnée $P = I + \varepsilon R(I, \varphi)$ de telle sorte que la P-composante de la solution cesse d'osciller.

Pour ce faire il faut éliminer les termes de l'ordre de ε dépendant de φ du second membre de l'équation pour $\dot{P}$.

En d'autres termes, il nous faut construire un difféomorphisme de $M: (I, \varphi) \mapsto (P, \varphi)$ tel que le champ perturbé soit envoyé dans un champ possédant sur chaque fibre une projection presque constante (à ε^2 près) sur la base.

En dérivant $P = I + \varepsilon h(I, \varphi)$ par rapport au temps et en groupant les termes du premier ordre en ε, on obtient

$$\dot{P} = \varepsilon [g + \omega\, \partial h/\partial\varphi] + r,$$

où ε l'argument de la fonction g a été remplacé par zéro; le résidu r (nous allons le voir plus bas) est une quantité du second ordre de petitesse par rapport à ε.

Nous allons essayer de choisir h de sorte à éliminer les termes du premier ordre en ε, c'est-à-dire de manière à annuler les crochets. On obtient formellement

$$h(I, \varphi) = -\frac{1}{\omega(I)} \int_{\varphi_0}^{\varphi} g(I, \psi, 0)\, d\psi.$$

(On se sert de l'hypothèse du théorème $\omega \neq 0$.) En réalité cette méthode de résolution de l'équation $g + \omega\partial h/\partial\varphi = 0$ est illicite, car la fonction h doit être 2π-périodique en φ pour que l'application $(I, \varphi) \mapsto (P, \varphi)$ soit définie sur M.

La formule précédente définit une fonction h sur un cercle (et non sur la droite qui est son revêtement universel) si seulement la valeur moyenne de la fonction g est nulle sur le cercle.

Donc, par le choix de h on peut éliminer non pas toute la perturbation g mais uniquement le terme oscillant

$$\tilde{g}(I, \varphi, 0) = g(I, \varphi, 0) - G(I).$$

La moyenne de la fonction $\tilde{g}$ par rapport à la période est maintenant égale à zéro et nous pouvons définir la fonction périodique h par la formule

$$h(I, \varphi) = -\frac{1}{\omega(I)} \int_{0}^{\varphi} \tilde{g}(I, \psi, 0)\, d\psi. \tag{1}$$

On obtient désormais l'équation en P

$$\dot{P} = \varepsilon G(P) + \varepsilon R.$$

Cette équation diffère de l'équation moyennisée

$$\dot{J} = \varepsilon G(J),$$

de la quantité εR de l'ordre de ε^2. Donc ses solutions s'écartent de celles de l'équation moyennisée à une vitesse de l'ordre de ε^2 et, par suite, d'une distance de l'ordre de ε pendant le temps $1/\varepsilon$. Comme l'écart entre P et I est aussi de l'ordre de ε, l'écart entre $I(t)$ et $J(t)$ reste de l'ordre de ε pendant une durée de temps de l'ordre de $1/\varepsilon$.

La démonstration de ce fait exige encore quelques évaluations simples des termes négligés.

C. Estimations.

1°. *Notations.* Soit $K \subset B$ un compact convexe contenant le point I_0. On suppose que $J(t)$ reste dans K pendant un temps T/ε. Notons $|\cdot|_0$ et $|\cdot|_1$ les normes dans les espaces C^0 et C^1 (le maximum du module d'une fonction et le maximum des modules d'une fonction et de ses dérivées premières). Soit c_1 une constante telle que

$$|f|_1 \leqslant c_1, \quad |g|_1 \leqslant c_1, \quad |\omega^{-1}|_1 \leqslant c_1 \text{ pour } I \in K.$$

2°. Prouvons que l'*application* $A: (I, \varphi) \longmapsto (P, \varphi)$ est un *difféomorphisme de* $K \times S^1$ *pour* ε *assez petits.*

◀ La définition de h (formule 1) entraîne $h \in C^1$. Donc $|\varepsilon h|_1 < 1$ pour ε assez petits. Si deux points avaient la même image par l'application A la différence des valeurs de εh en ces points serait égale à celle des valeurs de I; or ceci contredit l'inégalité $|\varepsilon h|_1 < 1$, puisque K est convexe. De $|\varepsilon h|_1 < 1$ il résulte également que A est un difféomorphisme local. Donc, A est un difféomorphisme. ▶

3°. *Majoration de* R. On a

$$R(P(I, \varphi, \varepsilon), \varphi, \varepsilon) = R_1 + R_2 + R_3 + R_4 + R_5,$$

$$R_1 = g(I, \varphi, 0) - g(P(I, \varphi, \varepsilon), \varphi, 0), \quad R_2 = g(I, \varphi, \varepsilon) - g(I, \varphi, 0),$$

$$R_3 = h(I, \varphi) - h(P(I, \varphi, \varepsilon), \varphi), \quad R_4 = \varepsilon g(I, \varphi, \varepsilon)\, \partial h/\partial I,$$

$$R_5 = \varepsilon f(I, \varphi, \varepsilon)\, \partial h/\partial \varphi.$$

Supposons que I et $P(I, \varphi, \varepsilon)$ appartiennent à K. Alors, puisque

$$P = I + \varepsilon h(I, \varphi)$$

on a

$$|R_1| \leqslant \varepsilon |g|_1 |h|_0, \quad |R_2| \leqslant \varepsilon |g|_1, \quad |R_3| \leqslant \varepsilon |h|_1 |h|_0,$$

$$|R_4| \leqslant \varepsilon |h|_1 |g|_0, \quad |R_5| \leqslant \varepsilon |h|_1 |f|_0.$$

Les normes de f, g et h sont évaluées au moyen de c_1. En définitive, si I et $P(I, \varphi, \varepsilon)$ appartiennent à K, on a

$$|R(P(I, \varphi, \varepsilon), \varphi, \varepsilon)| \leqslant c_2 \varepsilon,$$

où $c_2(c_1) > 0$ est une constante ne dépendant pas de I, φ, ε.

4°. *Majoration de* $P(t) - J(t)$.

Si l'on désigne par le symbole prime la dérivée par rapport au temps lent $\tau = \varepsilon t$, on constate que P et J vérifient les relations

$$P' = G(P) + \varepsilon R(P, \varphi(t), \varepsilon)$$

$$J' = G(J).$$

Donc $P - J = Z$ réalise l'inégalité

$$|Z|' \leqslant a|Z| + b,$$

où $a = |G|_1$, $b = c_2 \varepsilon$ tant que P, I et J restent dans K.

Posons $|Z(0)| = c$. La résolution de l'équation $z' = az + b$ avec la condition initiale c nous donne la majoration

$$|Z(\tau)| \leqslant (c + b\tau)\, e^{a\tau}$$

tant que P, I et J restent dans K.

5°. *Fin de la démonstration du théorème du point A.* ◀ Posons $c_3 = |h|_0$. Il vient

$$|P(I, \varphi, \varepsilon) - I| \leqslant c_3\varepsilon.$$

D'autre part, la majoration précédente nous permet d'écrire

$$|P(t) - J(t)| \leqslant c_4\varepsilon, \quad c_4 = (c_3 + c_2 T)\, e^{aT}$$

avec $\varepsilon t \leqslant T$ tant que $I(t)$, $P(t) = P(I(t), \varphi(t), \varepsilon)$ et $J(t)$ restent dans K.

Désignons par ρ la distance de la trajectoire du mouvement moyennisé $\{J(t), \varepsilon t \leqslant T\}$ à la frontière de K. Si $(c_3 + c_4)\, \varepsilon < \rho$, alors en vertu des majorations précédentes, $I(t)$, $P(t)$ et $J(t)$ restent dans K pour $\varepsilon t \leqslant T$. Or pendant toute cette durée

$$|I(t) - J(t)| \leqslant |I(t) - P(t)| + |P(t) - J(t)| \leqslant c_3\varepsilon + c_4\varepsilon. \text{ ▶}$$

D. Exemple.

L'équation

$$\ddot{x} = -x + \varepsilon(1 - x^2)\,\dot{x}$$

est dite *équation de Van der Pol*. C'est l'équation d'un pendule dans laquelle on a ajouté un « frottement » non linéaire, positif pour des amplitudes élevées et négatif pour de petites.

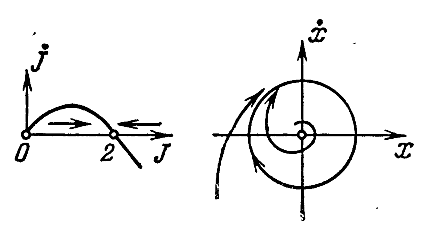


Fig. 94

L'équation non perturbée $\ddot{x} = -x$ se ramène à la forme standard $\dot{\varphi} = -1$, $\dot{I} = 0$ où $\varphi = \arg(x + i\dot{x})$, $2I = x^2 + \dot{x}^2$.

Dans le mouvement perturbé l'équation en I s'écrit

$$\dot{I} = \varepsilon(1 - x^2)\,\dot{x}^2 = 2\varepsilon I(1 - 2I\cos^2\varphi)\sin^2\varphi.$$

Donc, l'équation moyennisée s'écrit

$$\dot{J} = \varepsilon(J - J^2/2).$$

Cette équation possède une position d'équilibre répulsive $J = 0$ et une attractive $J = 2$.

Les positions d'équilibre de l'équation en J correspondent aux cycles du système perturbé. Le théorème prouvé plus haut nous per-

met d'affirmer que les variations de I dans le système perturbé sont proches de celles de J dans le système moyennisé pendant une durée de temps de l'ordre de $1/\varepsilon$. Mais si le système moyennisé possède une position d'équilibre non dégénérée (par exemple stable en première approximation), alors le système perturbé (pour ε assez petit) aura un cycle non dégénéré (par exemple, stable en première approximation) ; ceci est une conséquence simple du théorème des fonctions implicites.

En particulier, pour ε petit l'équation de Van der Pol possède un cycle limite stable, proche du cercle $x^2 + \dot{x}^2 = 4$ (fig. 94).

§ 18. Moyennisation dans les systèmes à plusieurs fréquences

Les systèmes à plusieurs fréquences sont bien moins étudiés que ceux à une fréquence. Ce paragraphe passe en revue les principaux résultats obtenus dans ce domaine.

A. Surfaces résonnantes.

Soit un système perturbé ordinaire de la méthode de moyennisation

$$\begin{cases} \dot{\varphi} = \omega(I) + \varepsilon f(I, \varphi, \varepsilon); & \varphi \in T^n, \ \varepsilon \ll 1, \ \omega \neq 0, \\ \dot{I} = \qquad \varepsilon g(I, \varphi, \varepsilon) & I \in B \subset \mathbb{R}^k. \end{cases}$$

Le vecteur des fréquences $\omega = (\omega_1, \ldots, \omega_n)$ est par définition résonnant si existe un vecteur non nul entier $m = (m_1, \ldots, m_n)$ tel que $(m, \omega) = 0$.

Le vecteur entier m est appelé *numéro de la résonance*.

Un point I de la base B est dit *résonnant* si le vecteur $\omega(I)$ est résonnant. Les points résonnants I correspondant à une même résonance de numéro m forment dans la base B du fibré une hypersurface

$$\Gamma_m = \{I \in B : (m, \omega(I)) = 0\}$$

appelée *hypersurface résonnante.*

Dans le cas général les points résonnants de même d'ailleurs que les points non résonnants sont partout denses dans B (si le nombre de fréquences $n > 1$).

E x e m p l e 1. Soit le système non perturbé à deux fréquences

$$\dot{\varphi}_1 = I_1, \quad \dot{\varphi}_2 = I_2, \quad \dot{I} = 0.$$

Ici B est le plan de coordonnées I_1, I_2 (privé du zéro, puisque nous avons admis que $\omega \neq 0$) ; les surfaces résonnantes sont les droites passant par 0 et dont la pente par rapport à I_1 est rationnelle.

Dans le cas d'un système à deux fréquences les surfaces résonnantes forment de même généralement une famille d'hypersurfaces ne se coupant pas (fig. 95 ; généralement = si le rang de $\partial\omega/\partial I$ est maximal). Dans ce cas la trajectoire du point I sur la base coupe généralement transversalement les surfaces résonnantes.

Tout autre est la disposition des surfaces résonnantes lorsque les fréquences sont au nombre de 3 ou plus.

E x e m p l e 2. Soit le système non perturbé à trois fréquences

$$\dot{\varphi}_1 = I_1, \quad \dot{\varphi}_2 = I_2, \quad \dot{\varphi}_3 = 1, \quad \dot{I} = 0.$$

Ici B est le plan rapporté aux coordonnées I_1, I_2 ; les surfaces résonnantes sont toutes les droites d'équations rationnelles.

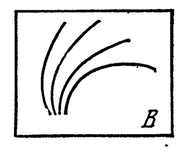
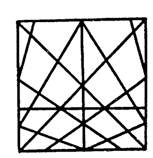


Fig. 95


Fig. 96

Dans ce cas la trajectoire générique du point I sur le plan coupe transversalement toutes les courbes résonnantes, mais beaucoup d'entre elles sous un petit angle, car tout voisinage aussi petit que l'on veut d'un élément linéaire contient un élément linéaire de courbe résonnante (fig. 96).

R e m a r q u e. Pour une meilleure intelligence de ce qui précède considérons l'application de la base dans un espace projectif à $(n - 1)$ dimensions

$$\Omega : B \to \mathbb{R}P^{n-1}, \quad \Omega(I) = (\omega_1(I) : \ldots : \omega_n(I)).$$

Les surfaces résonnantes sont les antécédents des hyperplans rationnels de $\mathbb{R}P^{n-1}$. Si $n = 2$, aux résonances correspondent des points rationnels sur la droite projective.

Si le nombre de fréquences $n > 2$, les hyperplans rationnels forment un ensemble connexe partout dense, de sorte que sur les résonances on peut se rendre d'un voisinage d'un point quelconque à un voisinage d'un autre point.

D'après ce qui précède, dans les systèmes à deux fréquences, le principal effet est le passage par les résonances; dans les systèmes à plus grand nombre de fréquences il faut nécessairement tenir compte des contacts entre les trajectoires de l'équation moyennisée et les surfaces résonnantes.

B. Influence d'une résonance.

Voyons sur des exemples simples l'effet éventuel d'une résonance.

E x e m p l e 1. Soit le système perturbé

$$\dot{\varphi}_1 = I_1, \quad \dot{\varphi}_2 = 1, \quad \dot{I}_1 = \varepsilon, \quad \dot{I}_2 = \varepsilon \cos \varphi_1.$$

Etudions la résonance $\omega_1 = 0$. Le mouvement moyennisé traverse la courbe résonnante $I_1 = 0$ à une vitesse non nulle. On calcule sans peine que dans la solution exacte la variation de I_2 entre $-\infty$ et $+\infty$ est donnée par l'intégrale de Fresnel

$$\Delta I_2 = \varepsilon \int_{-\infty}^{\infty} \cos (\varphi_0 + \varepsilon t^2/2)\, dt = c\, (\varphi_0) \sqrt{\varepsilon}.$$

Dans le système moyennisé J_2 est une intégrale première.

On remarquera que la principale contribution à l'intégrale est apportée par un voisinage de la résonance d'une largeur de l'ordre de $\sqrt{\varepsilon}$; l'intégrale elle-même est de l'ordre de $\sqrt{\varepsilon}$ et dépend de la phase initiale φ_0.

Donc, dans ce cas simple, l'intersection de la résonance entraîne une dispersion des solutions de l'équation perturbée prenant la même valeur initiale I d'une distance de l'ordre de $\sqrt{\varepsilon}$ l'une par rapport à l'autre, et, de plus, cette dispersion a lieu dans un voisinage de la surface résonnante d'une largeur de l'ordre de $\sqrt{\varepsilon}$.

La présence de quantités de l'ordre de $\sqrt{\varepsilon}$ est caractéristique de tous les problèmes comportant un passage par la résonance.

Si dans le premier exemple le passage par la résonance n'entraîne qu'une légère dispersion des projections des trajectoires du système perturbé sur la base par rapport aux trajectoires du système moyennisé, dans l'exemple qui va suivre les mouvements perturbé et moyennisé sont fondamentalement différents.

E x e m p l e 2. Soit le système perturbé

$$\dot{\varphi}_1 = I_1, \quad \dot{\varphi}_2 = I_2, \quad \dot{I}_1 = \varepsilon, \quad \dot{I}_2 = \varepsilon \cos (\varphi_1 - \varphi_2).$$

Le système moyennisé s'écrit:

$$\dot{J}_1 = \varepsilon, \qquad \dot{J}_2 = 0.$$

La solution du système moyennisé vérifiant la condition initiale $J_1(0) = 1$, $J_2(0) = 1$ pour $t = 1/\varepsilon$ est $J_1(t) = 2$, $J_2(t) = 1$.

La solution du système perturbé vérifiant la condition initiale $I_1(0) = 1$, $I_2(0) = 1$, $\varphi_1(0) = 0$, $\varphi_2(0) = 0$ est, pour $t = 1/\varepsilon$, $I_1(t) = 2$, $I_2(t) = 2$.

Donc, la projection du mouvement perturbé sur la base se déplace systématiquement dans une direction différente de celle de la

trajectoire du mouvement moyennisé. Pendant le temps $t = 1/\varepsilon$ ces deux trajectoires s'éloignent l'une de l'autre d'une grande distance (de l'ordre de 1).

L'équation moyennisée est impropre à la description du mouvement perturbé considéré, car ce dernier a lieu tout le temps sur une surface résonnante, or au voisinage de cette surface la moyennisation est illicite, car la moyenne temporelle n'est pas voisine de la moyenne spatiale sur le tore à n dimensions tout entier. L'attraction d'une

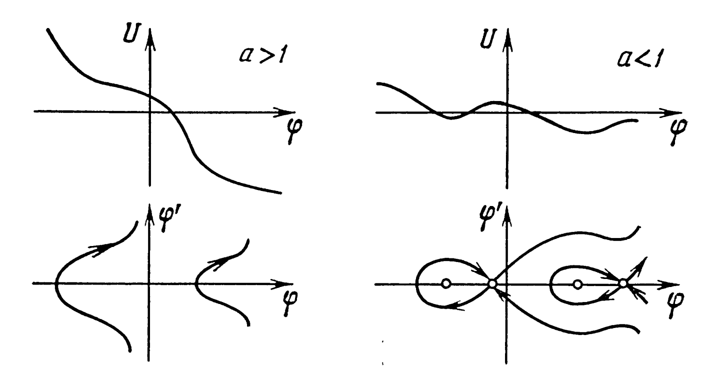


Fig. 97

partie des trajectoires par les mouvements résonnants est un phénomène générique pour les systèmes à plusieurs fréquences.

E x e m p l e 3. (A. Neichtadt). Soit le système *)

$$\dot{\varphi}_1 = I, \quad \dot{\varphi}_2 = 1, \quad \dot{I} = \varepsilon(a + \sin\varphi_1 - I).$$

Pour étudier ce système nous allons considérer l'équation d'un pendule avec moment de torsion et frottement, $\ddot{\varphi} = \varepsilon\,(a + \sin\varphi - \dot{\varphi})$ à laquelle ce système se ramène sans peine. Soit le temps lent $\tau = \sqrt{\varepsilon}\, t$ (à l'intervalle $t \sim 1/\varepsilon$ correspond l'intervalle $\tau \sim 1/\sqrt{\varepsilon}$). Si l'on désigne par le symbole prime les dérivées par rapport au temps lent on obtient l'équation

$$\varphi'' = a + \sin\varphi - \sqrt{\varepsilon}\,\varphi'.$$

Pour $\varepsilon = 0$ les solutions sont représentées sur la figure 97 (U est l'énergie potentielle).

On supposera que $a > 0$. On distinguera deux cas selon la valeur du moment de torsion a.

*) Obtenu par adjonction de l'équation triviale $\dot{\varphi}_2 = 1$ au système à une fréquence; le système obtenu est résonnant pour les valeurs nulles de la fréquence du système à une fréquence.

Si $a > 1$ (le moment de torsion est grand en regard du moment oscillant), le terme $\sin \varphi$ n'est pas essentiel : I varie de façon monotone. Le passage par la résonance $I = 0$ se traduit par un changement du sens de rotation du pendule.

Si $a < 1$ il est possible que le mouvement du pendule présente un régime oscillant (une boucle à l'intérieur de la séparatrice). Ce régime oscillant correspond aux trajectoires qui restent constamment au voisinage de la résonance.

Le petit frottement $\sqrt{\varepsilon}\varphi'$ a pour effet de détruire la boucle de l'intérieur de la séparatrice. A la place de l'intérieur de cette boucle

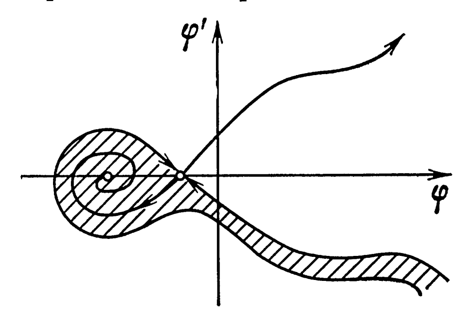


Fig. 98

on voit apparaître sur le plan (φ, φ') une bande étroite (d'une largeur de l'ordre de $\sqrt{\varepsilon}$) le long de la partie infinie de la séparatrice, constituée des points capturés par la position d'équilibre attractive; la région intérieure à la séparatrice est également capturée (fig. 98).

Si l'on revient au système initial, on s'aperçoit que pour $a < 1$ la résonance capture une petite portion de toutes les trajectoires (la mesure de l'ensemble des conditions initiales (I, φ) capturées pendant une durée de temps de $1/\varepsilon$ est de l'ordre de $\sqrt{\varepsilon}$). Pour les conditions initiales capturées, l'écart entre la variation de la variable lente I et celle de la solution de l'équation moyennisée J pendant une durée de temps de $1/\varepsilon$ est une quantité de l'ordre de 1.

Pour les autres conditions initiales (c'est-à-dire pour toutes les conditions initiales à l'exception de l'ensemble de mesure de l'ordre de $\sqrt{\varepsilon}$) cet écart reste petit (de l'ordre de $\sqrt{\varepsilon} \ln \varepsilon$ comme le prouve le calcul).

Si $a > 1$ il n'y a pas de capture par la résonance.

C. Passage par les résonances dans un système à deux fréquences. Soit le système à deux fréquences $\omega_1(I)$, $\omega_2(I)$:

$$\dot{\varphi}_1 = \omega_1 + \varepsilon f_1, \qquad \dot{\varphi}_2 = \omega_2 + \varepsilon f_2,$$
$$\dot{I}_1 = \varepsilon g_1, \qquad \dot{I}_2 = \varepsilon g_2.$$

Définition. On dit que ce système *vérifie une condition A* si la vitesse de variation du rapport des fréquences ω_1/ω_2 ne s'annule en aucun point des trajectoires du système perturbé:

$$\omega_2 \frac{\partial \omega_1}{\partial I} g \neq \omega_1 \frac{\partial \omega_2}{\partial I} g.$$

Ce système *vérifie la condition* $\overline{A}$ si la vitesse de variation du rapport des fréquences ω_1/ω_2 ne s'annule en aucun point des trajectoires du système moyennisé :

$$\omega_2 \frac{\partial \omega_1}{\partial I} G \neq \omega_1 \frac{\partial \omega_2}{\partial I} G.$$

On supposera que tous les systèmes envisagés sont analytiques.

Théorème. *Si un système vérifie la condition A, l'écart entre le mouvement lent I (t) dans le système perturbé et J (t) dans le système moyennisé reste petit pendant une durée de temps* $t = 1/\varepsilon$:

$$| I(t) - J(t) | \leqslant c \sqrt{\varepsilon} \text{ si } I(0) = J(0), \ 0 \leqslant t \leqslant 1/\varepsilon.$$

◀ Pour prouver ce théorème on considère un nombre fini de surfaces résonnantes de numéros peu élevés (grand pour les petits ε) et on effectue des changements ordinaires de variables (cf. § 17) en dehors de petits voisinages de ces surfaces.

Aux voisinages de ces surfaces on constate une dispersion de l'ordre de $\sqrt{\varepsilon}$ (comme dans les exemples précédents).

En faisant la somme de ces dispersions et des dérives qui ont lieu dans les intervalles des résonances on obtient la majoration du théorème. ▶

Pour les détails voir V. Arnold, Conditions d'applicabilité et estimation de la méthode de moyennisation pour les systèmes qui durant leur évolution passent par des résonances, DAN SSSR, 161, 1, 1965; A. Neichtadt, Sur le passage par des résonances dans un problème à deux fréquences, DAN SSSR, 22, 2, 1975; la thèse de A. Neichtadt « Sur quelques problèmes de résonance dans les systèmes non linéaires », MGU, 1975, contient la démonstration de la majoration $c \sqrt{\varepsilon}$ au lieu de la majoration $\sqrt{\varepsilon} \ln^2 \varepsilon$ qui figurait dans le premier des travaux cités.

Théorème (A. Neichtadt). *Si un système vérifie la condition* $\overline{A}$ *et une condition B (presque toujours réalisée), alors pour tous les points initiaux* (I_0, φ_0) *à l'exception d'un ensemble de mesure* $\leqslant c_1 \sqrt{\varepsilon}$, *l'écart entre le mouvement lent I (t) du système perturbé et le mouvement J (t) du système moyennisé reste petit pendant une durée de temps de* $1/\varepsilon$:

$$| I(t) - J(t) | \leqslant c_2 \sqrt{\varepsilon} \, | \ln \varepsilon | \text{ si } I(0) = J(0).$$

◀ Pour prouver ce théorème on considère un nombre fini de surfaces résonnantes de numéros peu élevés et on effectue un habituel changement de variables en dehors de petits voisinages de ces surfaces.

Pour étudier ces résonances on procède à une moyennisation sur des cercles qui sont trajectoires du mouvement non perturbé dans le cas de résonance.

Fixons à cet effet un numéro de résonance (m_1, m_2), où m_1 et m_2 sont premiers entre eux et remplaçons sur le tore les coordonnées angulaires (φ_1, φ_2) par les nouvelles coordonnées angulaires (γ, δ), où $\gamma = m_1\varphi_1 + m_2\varphi_2$. La vitesse de variation de la coordonnée angulaire γ est nulle dans le mouvement non perturbé en cas de résonance, car $m_1\omega_1 + m_2\omega_2 = 0$.

Sur la base on considère également la coordonnée spéciale $\rho = m_1\omega_1 + m_2\omega_2$. L'équation de la surface résonnante s'écrit maintenant $\rho = 0$, si bien que la quantité ρ caractérise la distance à la surface résonnante. On

désignera par σ un point de la surface résonnante. Au voisinage de cette surface on peut caractériser un point de la base par la distance ρ à la résonance et la projection σ sur la surface résonnante.

Dans les nouvelles coordonnées le système perturbé s'écrit

$$\dot{\gamma}=\rho+\varepsilon F_1, \quad \dot{\delta}=\alpha(I)+\varepsilon F_2, \quad \dot{\rho}=\varepsilon F_3, \quad \dot{\sigma}=\varepsilon F_4,$$

les fonctions F_k sont 2π-périodiques en γ et δ.

La moyennisation sur les trajectoires du mouvement résonnant se ramène à une moyennisation sur δ. Le système moyennisé s'écrit

$$\dot{\gamma}=\rho+\varepsilon G_1, \quad \dot{\rho}=\varepsilon G_3, \quad \dot{\sigma}=\varepsilon G_4.$$

Les fonctions G_k sont 2π-périodiques en γ et dépendent également de ρ et de ε.

Introduisons le temps lent $\tau = \sqrt{\varepsilon}\, t$ et la distance normée à la surface résonnante $r = \rho/\sqrt{\varepsilon}$. Si l'on désigne la dérivée par rapport à τ par le symbole prime, l'équation moyennisée devient

$$\gamma' = r + \sqrt{\varepsilon}\, G_1, \quad r' = G_3, \quad \sigma' = \sqrt{\varepsilon}\, G_4.$$

Les fonctions G_k dépendent de γ, $\sqrt{\varepsilon}\, r$, σ et ε.

En faisant $\varepsilon = 0$ dans cette équation, on obtient pour première approximation l'équation

$$\gamma' = r, \quad r' = u(\gamma, \sigma), \quad \sigma' = 0.$$

Donc, en première approximation, on obtient une équation du pendule à moment de torsion, dépendant du paramètre σ. L'hamiltonicité de cette équation, qui a été mise en évidence par le calcul, est un fait étonnant que rien ne laissait entrevoir.

Les trajectoires de cette équation ont la même allure sur le plan (γ, r) que celles de l'exemple 3 du point B pour $a < 1$ ou $a > 1$ selon que la fonction u change de signe ou non.

Les séparatrices ne présentent des boucles que pour un nombre peu élevé de résonances possédant des numéros pas très grand (on se sert ici de la condition $\bar{A}$). En effet, la condition $\bar{A}$ entraîne que la valeur moyenne de la fonction u par rapport à γ est non nulle. Pour les résonances de numéros élevés, la fonction u diffère peu de sa valeur moyenne (puisque la moyennisation sur δ est proche alors de la moyennisation sur le tore) et, par suite, elle est partout non nulle. Ce qui correspond à un pendule à moment de torsion élevé en regard du moment de poids. Dans ce cas l'équation de première approximation ne possède ni position d'équilibre ni domaine d'oscillations.

Lorsqu'on passe de l'équation de première approximation à l'équation complète, la boucle de la séparatrice se transforme en zone de capture de la résonance comme dans l'exemple 3 du point B *). La mesure de l'ensemble des points capturés est de l'ordre de $\sqrt{\varepsilon}$ si toutes les positions d'équilibre de l'équation de première approximation sont simples (c'est-à-dire si les zéros de la fonction u sont simples: $u = 0 \Rightarrow \partial u/\partial\gamma \neq 0$). Ceci constitue précisément la condition B du théorème de Neichtadt. On remarquera que cette condition porte sur les équations de première approximation correspondant à un nombre fini de résonances (puisque sous la condition $\bar{A}$ les équations de deuxième approximation ne possèdent des positions d'équilibre que pour un nombre fini de résonances).

---

*) Contrairement à l'exemple du point B, dans le cas général les trajectoires « capturées » ne sont pas tenues à demeurer pour toujours au voisinage de la résonance.

La démonstration du théorème s'achève par une addition des majorations des variations de I dans les intervalles des résonances et au voisinage des résonances dans la partie non capturée de l'espace des phases. Pour les détails on se reportera à la thèse citée plus haut. ▶

Remarque. Dans les systèmes à deux fréquences nous n'avons pas examiné les cas où la condition $\overline{A}$ est violée, c'est-à-dire celui où le rapport des fréquences du mouvement rapide varie de façon non monotone dans le mouvement moyennisé. Un tel comportement est impossible dans le cas d'une base à une dimension, mais si le nombre des variables lentes I est $\geqslant 2$ l'inversion de la vitesse de changement du rapport des fréquences est un phénomène générique qu'on ne peut éliminer par une déformation du système.

D. Systèmes à plusieurs fréquences.

Les systèmes à nombre de fréquences plus grand que deux sont bien moins étudiés que les systèmes à deux fréquences. Pour les systèmes génériques les fréquences du mouvement rapide sont incommensurables pour presque toutes les valeurs des variables lentes. Il est donc normal que pour la plupart des conditions initiales la méthode de moyennisation décrive correctement l'évolution des variables lentes sur des intervalles de temps de l'ordre de $1/\varepsilon$.

Dans ce domaine les premiers théorèmes généraux appartiennent à D. Anossov (D. Anossov, Moyennisation dans les systèmes d'équations différentielles ordinaires à solutions rapidement oscillantes, Izv. AN SSSR, ser. matem., 24, 5 (1960)), 721-742, et à T. Casuga (T. Casuga, On the adiabatic theorem for the hamiltonian system of differential equations in the classical mechanics, I, II, III, Proc. Japan acad. 37, 7, 1961).

Le théorème d'Anossov dit que pour tout nombre positif ρ la mesure de l'ensemble des conditions initiales (d'un compact de l'espace des phases) telles que

$$\max_{0\leqslant t\leqslant 1/\varepsilon} |I(t)-J(t)| > \rho \quad \text{pour } I(0)=J(0)$$

tend vers zéro avec ε (comme toujours I est la projection du mouvement perturbé, J le mouvement moyennisé ; on suppose que les fréquences sont indépendantes au sens que le rang de la dérivée des fréquences par rapport à la variable lente $\partial\omega/\partial I$ est égal au nombre de variables rapides).

Ce théorème a été prouvé en réalité sous des hypothèses bien plus larges : la quasi-périodicité des mouvements rapides n'est pas forcément prescrite, on suppose seulement l'ergodicité du mouvement rapide sur presque tous les tores ; on suppose comme toujours que la solution J de l'équation moyennisée est prolongeable pendant l'intervalle de temps $1/\varepsilon$.

L'ensemble de mesure petite avec ε, où sont possibles de grands écarts par rapport au mouvement moyennisé durant le temps $1/\varepsilon$, correspond à toutes les trajectoires capturées par les résonances ou errant le long des surfaces résonnantes, en passant de l'une à l'autre,

ce qui est également possible si le nombre de fréquences est plus grand que deux.

Il y a intérêt à estimer la mesure de cet ensemble. Pour les systèmes à deux fréquences par exemple, Neichtadt a prouvé (voir point C) que $|I(t) - J(t)| \leqslant c_2 \sqrt{\varepsilon}\,|\ln \varepsilon|$ en dehors d'un ensemble de mesure $< c_1 \sqrt{\varepsilon}$ (sous de faibles conditions imposées au système).

Supposons que les fréquences sont indépendantes, c'est-à-dire que le rang $\partial\omega/\partial I$ est égal au nombre de fréquences.

Théorème (A. Neichtadt). *Dans le cas d'un système à fréquences indépendantes, en dehors d'un ensemble de mesure* $\varkappa$ *petite, l'écart*

$$\max_{0 \leqslant t \leqslant 1/\varepsilon} |I(t) - J(t)| \quad avec \quad I(0) = J(0)$$

de la méthode de moyennisation est majoré par $c_3 \sqrt{\varepsilon/\varkappa}$.

Formulation équivalente. Soit $E(\varepsilon, \rho)$ l'ensemble des conditions initiales d'un compact fixe, telles que l'écart soit de ρ pour la valeur indiquée de ε.

On a

$$\text{mes}\, E(\varepsilon, \rho) \leqslant c_4 \sqrt{\varepsilon}/\rho.$$

La démonstration figure dans l'article de A. Neichtadt, Sur la moyennisation dans les systèmes à plusieurs fréquences, II, DAN SSSR, 226, 6 (1976), 1295-1298. Cette démonstration utilise l'idée du travail de Casuga cité plus haut, savoir un changement de variables de la méthode de moyennisation modifiée (lissée) de telle sorte qu'il soit défini par des fonctions différentiables non seulement en dehors des voisinages des résonances, mais partout.

Le résultat de Neichtadt peut être interprété comme l'indépendance statistique des accroissements de l'écart entre I et J sur des intervalles ultérieurs de temps de longueur 1. En effet, l'accroissement $I - J$ durant un intervalle de temps de l'ordre de 1 est une quantité de l'ordre de ε, quant au nombre d'intervalles de longueur T, contenus dans l'intervalle $1/\varepsilon$, il est de l'ordre de $1/\varepsilon$. Si les accroissements sur chaque intervalle T étaient indépendants, l'accroissement pendant l'intervalle de temps $1/\varepsilon$ serait, en vertu des lois de la théorie des probabilités, proportionnel au produit de l'accroissement pendant l'intervalle T par la racine carrée du nombre d'essais, c'est-à-dire une quantité de l'ordre de $\varepsilon \sqrt{1/\varepsilon} = \sqrt{\varepsilon}$.

Le théorème de Neichtadt fournit un même ordre de grandeur pour l'accroissement, mais pas pour toutes les conditions initiales: il faut exclure l'ensemble des conditions initiales de mesure de l'ordre de $\sqrt{\varepsilon}$, sur lequel il y a capture par les résonances et d'importants écarts, ne s'inscrivant pas dans le schéma des accroissements indépendants.

L'idée de l'indépendance des accroissements de l'écart entre I et J peut vraisemblablement être bien mieux justifiée dans le cas où le mouvement rapide n'est pas quasi périodique mais un C-système. C'est du moins ce que suggère le théorème limite central pour des fonctions de l'espace des phases (Y. Sinaï, Théorème limite central pour flots géodésiques sur des variétés à courbure négative constante, DAN SSSR, **133**, 6 (1960), 1303-1306; M. Ratner, Théorème limite central pour C-flots sur des variétés à trois dimensions, DAN SSSR, **186**, 3 (1969), 519-521). Ce théorème justifie les notions relatives au cas spécial où le mouvement lent et le mouvement rapide ne dépendent pas des variables lentes:

$$\dot{I} = \varepsilon g(\varphi), \qquad \dot{\varphi} = \omega(\varphi).$$

Les considérations probabilistes deviennent particulièrement intéressantes si l'on considère le comportement du système sur des intervalles de temps grands par rapport à $1/\varepsilon$ (disons de l'ordre de $1/\varepsilon\sqrt{\varepsilon}$ ou $1/\varepsilon^2$). Si pendant le temps $1/\varepsilon$ il y a capture de la $\sqrt{\varepsilon}$-ième partie de toutes les trajectoires par une résonance, et si sur les intervalles de temps ultérieurs de longueur $1/\varepsilon$ il y a capture de nouvelles trajectoires, alors au bout d'un intervalle de temps de l'ordre de $1/\varepsilon\sqrt{\varepsilon}$ la plupart des trajectoires seront capturées par les résonances et au bout du temps $1/\varepsilon^2$ on n'aura que des mouvements résonnants. Certes l'indépendance des captures sur des intervalles de temps consécutifs de longueur $1/\varepsilon$ est une hypothèse subsidiaire forte, et en plus de la capture par une résonance on observe le phénomène inverse.

Le champ d'action du théorème de Neichtadt est fortement restreint par l'indépendance des fréquences. La condition

le rang de $\partial\omega/\partial I$ est égal au nombre de fréquences

peut être remplacée par l'indépendance des rapports des fréquences:

le rang de l'application $I \longmapsto (\omega_1(I) : \ldots : \omega_n(I))$ est égal à $n - 1$.

Mais dans le cas où le nombre de variables lentes est petit (inférieur de plus d'une unité au nombre de fréquences) même cette condition n'est pas remplie.

La généralisation du théorème de Neichtadt au cas où le nombre de variables lentes est bien plus petit que le nombre de fréquences implique l'étude d'approximations diophantiennes sur des sous-variétés de l'espace euclidien $\mathbb{R}^n$.

Les applications

$$\omega: \mathbb{R}^k \to \mathbb{R}^n, \quad k < n,$$

vérifiant les conditions de non-dégénérescence (non-nullité de certains déterminants) semblent admettre la même minoration

$$|(m, \omega(I))| \geqslant C\,|m|^{-\nu}, \qquad m \in \mathbb{Z}^n \setminus 0$$

pour presque tous les I de $\mathbb{R}^k$ que pour presque tous les points de $\mathbb{R}^n$.

Des résultats analogues ont été obtenus pour des courbes spéciales ($\omega_s = I^s$); voir l'ouvrage de V. Sprindjouk, Problème de Maler dans la théorie métrique des nombres, Minsk, 1967; pour le cas général on peut se référer au travail de A. Partley, Approximations diophantiennes sur des sous-variétés d'un espace euclidien, Analyse fonctionnelle 3, 4 (1969), 59-62.

A signaler que ces travaux ne traitent ni de la généralisation du théorème de Neichtadt ni de l'estimation exacte de ν (ceci est du reste sans grande importance pour notre problème dans lequel un changement de la valeur de ν n'entraînera qu'un changement dans l'ordre des dérivées (dont on présuppose l'existence) des seconds membres des équations).

§ 19. Moyennisation dans les systèmes hamiltoniens

Dans ce paragraphe on décrit succinctement les particularités de la moyennisation dans le cas où système non perturbé et système perturbé sont hamiltoniens.

A. Calcul du système moyennisé.

Supposons que dans un système non perturbé on ait introduit des variables action-angle, c'est-à-dire des variables canoniquement conjuguées *) ($I_1, \ldots, I_n$; $\varphi_1, \ldots, \varphi_n \bmod 2\pi$) telles que le hamiltonien non perturbé H_0 dépende uniquement des variables d'action I. Les équations canoniques de Hamilton s'écrivent

$$\dot{\varphi} = \partial H/\partial I, \quad \dot{I} = -\partial H/\partial \varphi,$$

c'est-à-dire pour $H = H_0(I)$

$$\dot{\varphi} = \omega(I), \qquad \dot{I} = 0,$$

où le vecteur des fréquences $\omega(I)$ est égal à $\partial H_0/\partial I$.

Le système perturbé est défini par le hamiltonien $H = H_0(I) + \varepsilon H_1(I, \varphi, \varepsilon)$, où la fonction H_1 est 2π-périodique en les coordonnées angulaires φ. Donc, les équations du mouvement perturbé sont de la forme

$$\dot{\varphi} = \omega(I) + \varepsilon \partial H_1/\partial I, \quad \dot{I} = -\varepsilon \partial H_1/\partial \varphi.$$

Théorème. *Dans un système hamiltonien à n degrés de liberté et à n fréquences l'évolution des variables lentes n'a pas lieu au sens que le système moyennisé est de la forme* $\dot{J} = 0$.

◀ Pour calculer l'intégrale de $\partial H_1/\partial \varphi_s$ étendue à un tore à n dimensions on peut d'abord intégrer par rapport à la variable φ_s. Cette intégrale simple est égale à l'accroissement de la fonction périodique H_1 sur une période, c'est-à-dire à zéro. ▶

Ce théorème simple montre que l'évolution des variables lentes dans un système hamiltonien diffère fondamentalement des phénomènes observés dans les systèmes génériques (non hamiltoniens).

B. Théorème de Kolmogorov.

Supposons que les fréquences sont indépendantes au sens que la dérivée des fréquences par rapport aux variables d'action $\partial\omega/\partial I$ n'est pas dégénérée. Dans ce cas, selon A. Kolmogorov (A. Kolmogorov, Sur la conservation des mouvements quasi périodiques par une petite variation de la fonction de Hamilton, DAN SSSR, 98,

*) Les coordonnées (I, φ) sont *canoniquement conjuguées* si la structure symplectique de l'espace des phases s'écrit $\omega = \sum dI_k \wedge d\varphi_k$.

4 (1954), 527-530) la plus grande partie des tores invariants $I = \text{const}$ ne fait que se déformer et ne disparaît pas par une petite perturbation du hamiltonien : pour la plupart des conditions initales les trajectoires du système perturbé et du système non perturbé sont partout denses dans les tores invariants.

Si le jacobien de l'application de la surface à $(n-1)$ dimensions $H_0(I) = h$ dans un espace projectif à $(n-1)$ dimensions, défini par $I \mapsto (\partial H_0/\partial I_1 : \ldots : \partial H_0/\partial I_n)$, alors les tores invariants du système perturbé sont, à un reste de petite mesure près, partout denses dans $H(I, \varphi) = h$, variété de niveau du hamiltonien à $2n-1$ dimensions.

En particulier, si le nombre de fréquences $n = 2$, ces tores à deux dimensions partitionnent la variété de niveau à trois dimensions. Donc les variables d'action varient peu sur un intervalle de temps infini même pour les trajectoires qui ne sont pas situées sur les tores : toute trajectoire issue de l'intervalle compris entre deux tores invariants restera dans cet intervalle.

Si le nombre de fréquences est supérieur à deux, les tores ne partitionnent plus la variété de niveau du hamiltonien et certaines trajectoires (celles-ci constituent un ensemble de mesure petite) peuvent errer au voisinage des surfaces résonnantes entre les tores invariants et s'éloigner des valeurs initiales des variables d'action.

Il existe des exemples (V. Arnold, Sur l'instabilité des systèmes dynamiques à plusieurs degrés de liberté, DAN SSSR, **156**, 1 (1964), 9-12) dans lesquels cet éloignement a réellement lieu avec une vitesse moyenne exponentiellement petite (de l'ordre de $e^{-1/\sqrt{\varepsilon}}$).

C. Théorème de Nekhorochev.

La vitesse moyenne à laquelle les variables d'action s'éloignent de leurs valeurs initiales dans des systèmes hamiltoniens génériques quelconques est si petite qu'elle n'est localisée par aucune approximation de la théorie des perturbations (c'est-à-dire cet éloignement n'est pas perceptible au bout d'un temps de l'ordre de $1/\varepsilon^N$ quel que soit N, ε étant le paramètre de la perturbation).

Plus exactement, N. Nekhorochev (N. Nekhorochev, Sur le comportement de systèmes hamiltoniens proches de systèmes intégrables, Analyse fonctionnelle et ses applications, 5, 4 (1971) ; N. Nekhorochev, Estimation exponentielle de la durée de stabilité de systèmes hamiltoniens proches de systèmes intégrables, UMN, **32**, 6 (1977), (5-66) ; voir également sa thèse, MGU, 1973) a montré que pour presque tout hamiltonien non perturbé $H_0(I)$ il existe des nombres positifs a et b tels que la vitesse moyenne de variation des variables d'action I dans le système perturbé n'est pas supérieure à ε^b pendant une durée de temps $T = e^{1/\varepsilon^a}$.

A noter que T croît plus vite que toute puissance de $1/\varepsilon$ pour $\varepsilon \to 0$, de sorte que I varie peu pendant le temps $1/\varepsilon^N$, $\forall N$.

Les constantes a et b dépendent des propriétés géométriques du hamiltonien non perturbé H_0. Par exemple, si la fonction H_0 est strictement convexe (la matrice de $\partial^2 H_0/\partial I^2$ est définie positive), on peut prendre $a = 2/(6n^2 - 3n + 14)$, $b = 3a/2$, où n est le nombre de fréquences.

Le théorème est prouvé pour presque tous les H_0: ne sont exclues que les fonctions H_0, solutions d'une infinité d'équations algébriques en les coefficients de Taylor. N. Nekhorochev appelle ces fonctions exclusives fonctions *non raides*. Pour les fonctions H_0 non raides l'éloignement est possible pendant une durée de temps de l'ordre de $1/\varepsilon$. Dans les exemples d'éloignement exponentiellement lent (cf. point B) la fonction H_0 est raide.

La démonstration du théorème de Nekhorochev se base sur la propriété simple suivante de la moyennisation dans un système hamiltonien.

Supposons qu'un système hamiltonien à n fréquences présente les résonances $(m, \omega) = 0$ pour certaines valeurs des variables lentes I. On a donc intérêt au voisinage de la surface résonnante correspondante à effectuer la moyennisation non pas sur des tores à n dimensions, mais sur des tores résonnants de dimension inférieure. Le tore résonnant est de dimension $n - 1$ si la résonance est simple, c'est-à-dire si la direction du vecteur entier m est définie de façon unique. Si l'équation $(m, \omega) = 0$ en m possède k solutions rationnellement indépendantes, les trajectoires du mouvement rapide seront partout denses dans les tores résonnants de dimension $n - k$ et c'est sur ces tores qu'il faut effectuer la moyennisation.

Théorème. *Dans la moyennisation sur les tores résonnants correspondant aux résonances* $(m, \omega) = 0$, *la direction de l'évolution des variables d'action I du système moyennisé est contenue dans le plan construit sur les vecteurs résonnants m* *) (*dans le cas d'une résonance simple la direction de l'évolution est définie de façon unique* : *c'est celle du support du vecteur m*).

◀ Etudions par souci de simplicité le cas où la résonance est simple. Soit γ la coordonnée angulaire invariante par la résonance $\gamma = (m, \varphi)$. Pour moyenniser le système perturbé il suffit de moyenniser le hamiltonien sur les variables rapides. On obtient alors le hamiltonien $H_0 + \varepsilon \overline{H}_1$, où $\overline{H}_1$ dépend des variables d'action et d'une seule coordonnée angulaire: γ.

Les équations du mouvement moyennisé nous donnent maintenant

$$\dot{I} = \varepsilon \frac{\partial \overline{H}_1}{\partial \varphi}.$$

*) On remarquera que dans l'espace des variables d'action, la structure affine est définie de façon intrinsèque et que l'identification des vecteurs de l'espace dual à l'espace des fréquences, aux vecteurs de l'espace des variables d'action est aussi définie de façon intrinsèque.

Or $\frac{\partial \bar{H}_1}{\partial \varphi} = \left(\frac{\partial \bar{H}_1}{\partial \gamma}\right)\left(\frac{\partial \gamma}{\partial \varphi}\right)$ et le vecteur $\frac{\partial \gamma}{\partial \varphi} = m$ sont de même direction. ▶

Le théorème de Nekhorochev se déduit du théorème démontré pour les raisons suivantes. Une évolution rapide (à une vitesse de l'ordre de ε) n'est possible que dans le cas d'une résonance et uniquement dans les directions engendrées par les vecteurs résonnants. Or les conditions de raideur imposées à H_0 (par exemple une convexité stricte de H_0) nous garantissent que cette évolution a lieu dans un sens partant de la surface résonnante. Donc, la résonance est violée et l'évolution n'a lieu que pendant un court intervalle de temps et par suite sa vitesse moyenne est majorée par une quantité exponentiellement petite.

Si les conditions de raideur ne sont pas remplies, il est possible de trouver sur la surface résonnante une courbe dont toute tangente appartient à un plan construit sur des vecteurs résonnants. Sur une telle courbe l'évolution peut avoir lieu à une vitesse moyenne de l'ordre de ε, ce qui éloigne les variables d'action de leurs valeurs initiales pendant un intervalle de temps de l'ordre de $1/\varepsilon$.

§ 20. Invariants adiabatiques

On passe en revue les principaux résultats de la théorie des invariants adiabatiques dans des systèmes hamiltoniens de paramètres à variations lentes.

A. Notion d'invariant adiabatique.

Lorsqu'on étudie des systèmes hamiltoniens de paramètres à variation lente, on est confronté à un phénomène singulier : des grandeurs, en principe indépendantes, deviennent asymptotiquement (c'est-à-dire lorsque la vitesse de variation des paramètres tend vers zéro) fonctions l'une de l'autre.

Considérons, par exemple, le mouvement d'un pendule de longueur variable. La longueur du pendule et l'amplitude des oscillations sont en principe indépendantes : si la longueur du pendule change et qu'elle reprenne ensuite sa valeur initiale l'amplitude des oscillations varie en général en fonction du mode de variation de la longueur.

Cependant si la longueur du pendule change assez lentement l'amplitude des oscillations, lorsque la longueur reprend sa valeur initiale, ne varie pratiquement pas. Bien plus, le rapport de l'énergie du pendule à la fréquence gardera pratiquement la même valeur tout au long du processus, même si énergie et fréquence subissent des variations.

Les grandeurs qui conservent asymptotiquement leurs valeurs par une variation assez lente des paramètres d'un système hamiltonien sont dites invariants adiabatiques.

Plus exactement soit le système hamiltonien $\dot{x} = v(x, \lambda)$, λ est un paramètre.

Une fonction I de point x et de λ est par définition un *invariant adiabatique* si pour toute fonction différentiable $\lambda(\tau)$ du temps lent $\tau = \varepsilon t$ la grandeur $I(x(t), \lambda(\varepsilon t))$, où $x(t)$ est solution de l'équation $\dot{x} = v(x, \lambda(\varepsilon t))$, varie peu sur l'intervalle de temps $0 \leqslant t \leqslant 1/\varepsilon$ si ε est assez petit.

B. Construction d'un invariant adiabatique d'un système à un degré de liberté.

Supposons que le hamiltonien $H(p, q; \lambda)$ possède pour chaque valeur du paramètre λ des trajectoires fermées $H(p, q; \lambda) = h$ (par exemple, des trajectoires entourant la position d'équilibre en laquelle la fréquence des petites oscillations est non nulle).

Désignons par $I(p, q; \lambda)$ l'aire interceptée par la trajectoire passant par le point de coordonnées (p, q) à λ fixe, divisée (traditionnellement) par 2π. La quantité I est dite *variable d'action.*

E x e m p l e. Pour le pendule, $H = \frac{ap^2}{2} + \frac{bq^2}{2}$; la trajectoire $H = h$ est une ellipse d'aire $\pi\sqrt{2h/a}\,\sqrt{2h/b} = 2\pi h/\sqrt{ab}$. La fréquence des oscillations est $\omega = \sqrt{ab}$. Donc, pour le pendule

$$I = H/\omega.$$

Le paramètre λ est ici le couple (a, b).

T h é o r è m e. *La variable d'action I est invariant adiabatique d'un système hamiltonien à un degré de liberté.*

C. Démonstration de l'invariance adiabatique de l'action.

La démonstration est fondée sur la méthode de moyennisation. Soit φ une coordonnée angulaire sur des trajectoires fermées. La coordonnée φ est choisie de façon à varier proportionnellement au temps de parcours de la trajectoire et à croître de 2π à chaque tour (il est bien entendu que la coordonnée angulaire φ et la variable d'action I dépendent aussi bien des coordonnées (p, q) que du paramètre λ).

L'équation de notre système à λ fixé prend la forme du système non perturbé standard de la méthode de moyennisation:

$$\dot{\varphi} = \omega(I, \lambda(\tau)), \quad \dot{I} = 0, \quad \dot{\tau} = 0.$$

Si maintenant λ varie lentement on obtient le système perturbé

$$\dot{\varphi} = \omega + \varepsilon f, \quad \dot{I} = \varepsilon g, \quad \dot{\tau} = \varepsilon,$$

où les fonctions f et g sont 2π-périodiques en φ.

Composons le système moyennisé.

L e m m e. *La variable d'action est intégrale première du système moyennisé (c'est-à-dire la moyenne de g sur φ est nulle).*

◀ Considérons le domaine intercepté par la trajectoire fermée $I = I_0$ du système correspondant à la valeur initiale du paramètre. En vertu du théorème de moyennisation l'image de ce domaine au bout d'un temps t quelconque de l'intervalle $[0, 1/\varepsilon]$ est à ε près le domaine intercepté par une trajectoire fermée $I = I_t$ du système correspondant à $\lambda = \lambda(\varepsilon t)$.

Or les équations du mouvement sont hamiltoniennes (quoique non autonomes). Donc, en vertu du théorème de Liouville, l'aire de l'image est égale à celle de l'antécédent. D'où $I_t = I_0$. ▶

C o r o l l a i r e. *Le rapport de l'énergie du pendule à la fréquence est un invariant adiabatique.*

E x e r c i c e. Une bille est animée d'un mouvement horizontal entre deux parois verticales parfaitement élastiques dont l'écartement varie lentement. Montrer que le produit de la vitesse de la bille par la distance entre les parois est un invariant adiabatique.

E x e r c i c e. Une particule chargée est en mouvement dans un champ magnétique variant peu lorsque la particule parcourt une hélice de Larmore autour d'une ligne du champ. Montrer que le rapport du carré de la projection de la vitesse de la particule sur une normale à la ligne du champ magnétique, soit $v_\perp^2/H$ est un invariant adiabatique (voir par exemple L. A r t s i m o v i t c h, Réactions thermonucléaires contrôlées, M., Fizmatguiz, 1961).

D. Invariants adiabatiques dans les systèmes hamiltoniens à plusieurs fréquences.

Soit un système hamiltonien à plusieurs fréquences $\dot{p} = -H_q$, $\dot{q} = H_p$ dépendant d'un paramètre λ et admettant, à λ fixe, les variables action-angle: $\dot{\varphi} = \omega(I, \lambda)$, $\dot{I} = 0$ (où $\omega = \partial H_0/\partial I$) de hamiltonien $H_0(I, \lambda)$ dépendant des n variables d'action de façon non dégénéréé, de sorte que

$$\det(\partial\omega/\partial I) = \det(\partial^2 H_0/\partial I^2) \neq 0.$$

Supposons comme précédemment que le paramètre λ commence à varier lentement. Les variations de p et q sont régies par des équations de Hamilton de fonction variable H, et les variables I par un système perturbé (nous supposons que $\lambda = \varepsilon t$, où ε est un petit paramètre).

L e m m e. *Le système perturbé est hamiltonien et de fonction de Hamilton $H = H_0(I, \lambda) + \varepsilon H_1(I, \varphi, \lambda, \varepsilon)$, univalente.*

La démonstration de ce lemme implique soit une incursion dans la géométrie symplectique ou le formalisme hamiltonien (voir par exemple V. A r n o l d, Méthodes mathématiques de la mécanique classique, Traduction française, Editions Mir, 1976), soit des calculs pénibles que nous omettrons.

C o r o l l a i r e. *Les variables d'action I sont intégrales premières du système moyennisé.*

◀ En effet, la fonction moyennisée qui constitue le second membre de l'équation $\dot{I} = -\varepsilon\partial H_1/\partial\varphi$, est la dérivée d'une fonction périodique ; sa valeur moyenne est donc nulle (cf. théorème du point A, § 19). ▶

Ce corollaire et le théorème de Neichtadt (cf. point D, § 18) nous conduisent à la conclusion suivante :

Dans un système hamiltonien à plusieurs fréquences de paramètres à variation lente, la variation des variables d'action I reste inférieure à ρ pendant un intervalle de temps de $1/\varepsilon$ si l'on néglige l'ensemble des conditions initiales de mesure $\leqslant c\sqrt{\varepsilon}/\rho$ dans l'espace des phases initial (on suppose que l'espace des phases est un compact et que la dérivée $\partial\omega/\partial I$ n'est pas dégénérée).

D é f i n i t i o n. Une fonction F (dépendant d'un point et d'un paramètre) d'un système hamiltonien est *presque invariant adiabatique* si pour tout $\rho > 0$ la mesure de l'ensemble des conditions initiales telles que la variation de la fonction F le long d'une solution des équations de Hamilton de paramètre à variation lente soit supérieure à ρ pendant la durée de temps $1/\varepsilon$ tend vers zéro avec ε.

Donc, *les variables d'action $(I_1, \ldots, I_n)$ sont presque invariants adiabatiques d'un système hamiltonien non dégénéré à plusieurs fréquences.*

E. Comportement des invariants adiabatiques pour $t \gg 1/\varepsilon$.

Même si un invariant adiabatique varie peu pendant un intervalle de temps de $1/\varepsilon$, nous n'avons aucune raison de supposer que cette variation reste petite sur de plus grands intervalles de temps (de l'ordre de $1/\varepsilon^2$, par exemple) ou à fortiori sur un intervalle de temps infini.

E x e m p l e. Soit un pendule de paramètre à variation lente et périodique

$$\ddot{x} = -\omega^2 (1 + a\cos\varepsilon t)\, x$$

Pour des ε aussi petits que l'on veut (c'est-à-dire pour une variation aussi lente que l'on veut du paramètre) il existe des résonances paramétriques telles que la position d'équilibre $x = 0$ devienne instable. Il est clair que l'invariant adiabatique d'un pendule linéaire varie indéfiniment (durant un intervalle de temps infini) en cas de résonance paramétrique.

Un tel comportement de l'invariant adiabatique d'un système de paramètre à variation lente et périodique est précisément lié à la linéarité de ce système, ou plus exactement à l'indépendance de la période des oscillations par rapport à l'amplitude. Si dans un système hamiltonien de paramètre à variation lente et périodique la dérivée de la fréquence du mouvement rapide par rapport à la variable d'action est non nulle, alors la variable d'action varie peu sur un

intervalle de temps infini (voir V. A r n o l d, Sur le comportement d'un invariant adiabatique par une variation lente périodique du hamiltonien, DAN SSSR **142**, 4, 1962, 758-761).

La démonstration est basée sur l'existence de tores invariants comme dans le théorème de Kolmogorov (cf. point B, § 19).

Un autre cas intéressant est celui où le paramètre varie de façon à posséder des limites bien définies pour $t \to -\infty$ et pour $t \to +\infty$. Dans ce cas il est possible d'envisager la valeur de l'invariant adiabatique à $-\infty$, à $+\infty$, et son accroissement sur un intervalle de temps infini

$$\Delta I = I(+\infty) - I(-\infty).$$

S'agissant de l'équation linéaire

$$\ddot{x} = -\omega^2(\varepsilon t)\, x, \quad \omega(-\infty) = \omega_-, \quad \omega(+\infty) = \omega_+,$$

on démontre que l'accroissement de l'invariant adiabatique sur un intervalle de temps infini est une quantité exponentiellement petite en ε (sous l'hypothèse de l'analyticité de la fonction ω qui ne doit pas changer de signe et se conduire « raisonnablement » à l'infini). Bien plus, on peut exhiber explicitement le terme principal de l'asymptotique de l'accroissement de l'invariant adiabatique pour $\varepsilon \to 0$ (cf. A. Dykhné, Passages quantiques dans l'approximation adiabatique, JETF, 38, 2, 1960, 570-578). Des résultats analogues ont été obtenus pour les systèmes linéaires à plusieurs dimensions. On trouvera les formulations précitées et les démonstrations dans l'article de M. F é d o r i o u k, Invariant adiabatique d'un système d'oscillateurs linéaires et théorie de la dispersion, Equations différentielles, **12**, 6 (1976), 1012-1018 (les références aux travaux antérieurs de physique sont omises dans cet article).

Le problème de l'accroissement de l'invariant adiabatique d'un système non linéaire à une dimension a également été étudié par des physiciens: on a montré la petitesse de l'accroissement par rapport à ε^N, c'est-à-dire l'absence de variations de l'invariant adiabatique à $O(\varepsilon^N)$ près pour tout N. (A. L e n a r d, Ann. of Physics, **6** (1959), 261-276). A. Neichtadt a obtenu une estimation exponentielle dans le cas analytique.

S'agissant des systèmes linéaires à plusieurs degrés de liberté l'invariance adiabatique des variables d'action, contrairement aux affirmations des physiciens, n'a généralement pas lieu: ces quantités ne sont que presque invariants adiabatiques, c'est-à-dire varient peu pour la plupart des conditions initiales.

§ 21. Moyennisation dans le feuilletage de Seifert

Lorsqu'on étudie le voisinage d'une trajectoire fermée on rencontre des cas où les trajectoires proches se bouclent aussi en première approximation, mais après avoir effectué plusieurs tours le long de la trajectoire fermée initiale (cas résonnant). L'étude du système au voisinage d'un mouvement périodique résonnant ou proche d'une résonance conduit à une variante spéciale de la méthode de moyennisation : la moyennisation dans le feuilletage de Seifert.

A. Feuilletage de Seifert.

Le feuilletage de Seifert est une partition du produit cartésien $\mathbb{R}^2 \times S^1$ en cercles. Cette partition s'effectue comme suit. Soit dans l'espace euclidien $\mathbb{R}^3$ un cylindre de bases horizontales et d'axe vertical. Partageons l'intérieur du cylindre en segments verticaux. Faisons coïncider la base inférieure avec l'image de la base supérieure par une rotation de $2\pi p/q$ autour de l'axe (en d'autres termes, nous faisons coïncider le point $(z, 0)$ de la base inférieure avec le point $(Az, 1)$ de la base supérieure, A étant une rotation d'angle $2\pi p/q$, p et q des nombres premiers entre eux).

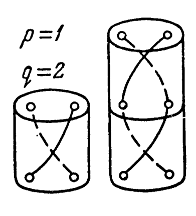


Fig. 99

D é f i n i t i o n. On appelle *feuilletage de Seifert* de type (p, q) une variété à trois dimensions $\mathbb{R}^2 \times S^1$ et sa partition en cercles suivant la procédure précédente.

Donc, tout cercle du feuilletage de Seifert s'obtient par collage de q segments à l'exception d'un seul, le cercle central qui provient de l'axe du cylindre.

Soit un revêtement à q feuillets de l'espace $\mathbb{R}^2 \times S^1$ d'un feuilletage de Seifert de type (p, q). L'espace de revêtement est lui-même difféomorphe à $\mathbb{R}^2 \times S^1$. Le feuilletage de Seifert dans la variété initiale induit une partition en cercles sur l'espace de revêtement. Cette partition peut être considérée comme un feuilletage de Seifert de type $(p, 1)$. (La rotation est maintenant d'angle $2\pi p$.)

Or le feuilletage de Seifert de type $(p, 1)$ est un fibré en cercles plus précisément un produit cartésien. Tous les cercles du feuilletage initial sont recouverts difféomorphement par q cercles fibres de ce fibré, sauf le cercle central qui, lui, est recouvert par une fibre à q feuillets (fig. 99).

B. Définition de la moyennisation dans le feuilletage de Seifert.

Soit donné un champ de vecteurs dans l'espace $\mathbb{R}^2 \times S^1$ du feuilletage de Seifert. Un champ de vecteurs est alors défini sur le fibré revêtement. Tout vecteur de ce champ se projette sur la base $\mathbb{R}^2$ du fibré revêtement. Prenons la moyenne du vecteur obtenu sur la

base le long d'une fibre du fibré revêtement. On obtient un vecteur bien défini en chaque point de la base. On a donc défini un champ de vecteurs sur la base. Cette opération s'appelle *moyennisation du champ initial le long d'un feuilletage de Seifert.*

En d'autres termes, la moyennisation sur un feuilletage de Seifert de type (p, q) se définit comme une moyennisation ordinaire sur son recouvrement à q feuillets.

C. Propriétés du champ moyennisé.

Si l'on procède à une moyennisation dans un fibré ordinaire on obtient un champ de vecteurs quelconque sur la base. Dans le cas d'un feuilletage de Seifert le champ de vecteurs est doué de propriétés spéciales: par exemple, le vecteur du champ moyennisé s'annule nécessairement au point central si $q > 1$.

Théorème. *Le champ de vecteurs obtenu par moyennisation sur un feuilletage de Seifert de type (p, q) est invariant par une rotation du plan d'angle $2\pi/q$.*

◀ Nous pouvons identifier la base du fibré avec l'une des bases du cylindre initial. La moyennisation sur le feuilletage de Seifert s'identifie alors avec la moyennisation sur q segments parallèles à l'axe du cylindre. Ces segments se correspondent par une rotation d'angle $2\pi/q$. Il est aisé de voir que la moyennisation commute avec la rotation d'angle $2\pi/q$ (après la rotation on doit prendre la moyenne sur les mêmes segments mais dans un ordre différent). ▶

D. Exemple.

Soit l'équation différentielle

$$\dot{z} = i\omega z + \varepsilon f(z, t), \quad z \in \mathbb{C},$$

f est une fonction complexe (pas forcément holomorphe) de période 2π en le temps réel t, ε un petit paramètre. L'équation correspondant à $\varepsilon = 0$ sera dite *non perturbée.*

Supposons que la fréquence ω du mouvement non perturbé est rationnelle ou proche d'un nombre rationnel p/q.

Les trajectoires de l'équation non perturbée avec $\omega = p/q$ forment dans $\mathbb{C} \times S^1 = \{z, t \bmod 2\pi\}$ un feuilletage de Seifert du type (p/q).

La moyennisation sur ce feuilletage nous donne l'équation moyennisée

$$\dot{z} = \varepsilon F(z),$$

où le champ de vecteurs F est associé à lui-même par une rotation d'angle $2\pi/q$ du plan de la variable z.

E. Coefficients de Taylor d'un champ symétrique.

Soit sur le plan de la variable complexe z un champ de vecteurs défini par une fonction complexe (pas forcément holomorphe) F. La

série de Taylor de la fonction complexe F en x et y ($z = x + iy$) s'écrit comme une série de Taylor sur les variables z et $\bar{z}$, soit

$$\sum F_{k,l} z^k \bar{z}^l .$$

P r o p o s i t i o n. *Si le champ F est invariant par une rotation d'angle $2\pi/q$, alors de tous les coefficients $F_{k,l}$ seuls sont non nuls ceux pour lesquels $k - l$ et 1 sont congrus modulo q.*

◀ La série de Taylor est unique. Donc chacun de ses termes définit un champ de vecteurs invariant par une rotation. Lorsque z subit une rotation de $2\pi/q$, le vecteur $z^k\bar{z}^l$ pivote d'un angle de $(k - l)\, 2\pi/q$. Cette rotation est une rotation d'angle $2\pi/q$ si et seulement si $k - l$ et 1 sont congrus modulo q. ▶

Soit le réseau des points entiers non négatifs (k, l). Repérons ceux pour lesquels $k - l$ et 1 sont congrus modulo q. Parmi ces points on aura toujours le point (1, 0) et tous les points entiers de la droite issue de (1, 0) et parallèle à la bissectrice du réseau. Ces points correspondent aux champs $z\Phi\,(|\,z\,|^2)$ invariants par une rotation quelconque.

Parmi les points repérés on aura toujours le point $(0, q - 1)$. Ce point correspond au champ $\bar{z}^{q-1}$ invariant par une rotation d'angle $2\pi/q$. Les points repérés engendrent un faisceau de demi-droites parallèles à la bissectrice et issues des points $(0, mq - 1)$ et $(mq + 1, 0)$.

F. Cas d'une symétrie d'ordre 3.

Soient des champs de vecteurs invariants par le groupe de symétries d'ordre 3 (c'est-à-dire on étudie le cas $q = 3$).

Les monômes de plus petit degré de la série de Taylor d'un champ symétrique par une rotation de 120° sont définis par les points repérés (k, l) du plan, tels que $k + l$ soit minimal. Les deux premiers monômes sont z et $\bar{z}^2$. Donc, tout champ du plan, invariant par une rotation d'angle $2\pi/3$, est de la forme

$$F(z) = az + b\bar{z}^2 + O(|\,z\,|^3).$$

Si l'on élimine le dernier terme on obtient une équation différentielle simple à symétrie d'ordre 3

$$\dot{z} = az + b\bar{z}^2.$$

les coefficients a, b et la coordonnée z sont complexes.

Supposons que $a \neq 0$, $b \neq 0$ et qu'en multipliant z par un nombre et en modifiant l'unité de temps, on puisse faire $b = 1$ et $|\,a\,| = 1$. La figure 100 montre comment varient les trajectoires de l'équation pour $a = e^{i\varphi}$, $b = 1$. Quel que soit a on a 4 positions d'équili-

bre aux sommets et au centre d'un triangle équilatéral. Pour a imaginaire pur, le système est hamiltonien. Pour étudier le système quel que soit a, il suffit de remarquer qu'il se déduit toujours à partir de ce système hamiltonien par une multiplication formelle des variables

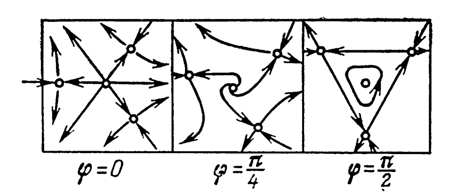


Fig. 100

z et t par des nombres complexes (c'est-à-dire par une similitude du plan z et une rotation du champ hamiltonien d'un angle constant).

G. Prise en considération des termes éliminés.

Essayons maintenant de tenir compte des termes éliminés $O(|z|^3)$. Supposons que $|a|$ est petit (ce qui correspond au voisinage d'une résonance d'ordre 3 dans le système d'équations différentielles initial). La dimension du triangle des points singuliers est petite (de l'ordre de $|a|$). Etudions notre champ de vecteurs symétriques au voisinage du point $z=0$, un voisinage grand par rapport à $|a|$, mais petit par rapport à 1.

Dans ce voisinage les termes $O(|z|^3)$ éliminés sont petits en comparaison des autres. D'où l'on déduit sans peine que leur prise en considération n'apporte pas de modification notable à l'allure des trajectoires si le système était structurellement stable. Dans le cas considéré le système est structurellement instable seulement pour les a imaginaires purs (le système est hamiltonien). L'hamiltonicité n'est pas conservée si l'on tient compte des termes éliminés.

Pour toute droite du plan de la variable complexe a, non confondue avec l'axe imaginaire, les trajectoires du système complet pour $|a| \neq 0$ assez petit (sous réserve que $b \neq 0$) sont, dans un voisinage de l'origine des coordonnées, petit par rapport à 1, et grand par rapport à a, tracées sur la figure 100, $\varphi \neq \pm \pi/2$.

La modification des trajectoires au passage du point a par l'axe imaginaire constitue un problème spécial sur lequel nous reviendrons au chapitre 6. Dans le cas générique ce changement est défini encore par un terme de la série de Taylor : tout se passe comme pour l'équation

$$\dot{z} = az + \bar{z}^2 + cz|z|^2,$$

où Re $c \neq 0$.

H. Application à l'équation initiale.

L'analyse de l'équation moyennisée nous fournit une riche information sur le système initial dans le cas où le paramètre ε est assez petit. Sans nous appesantir sur les justifications nous allons donner seulement la traduction des résultats obtenus en termes de trajectoires de l'équation initiale.

Les trois positions d'équilibre qui occupent les sommets d'un triangle équilatéral correspondent à une courbe intégrale fermée de l'équation initiale. Lorsque l'écart entre la fréquence ω du mouve-

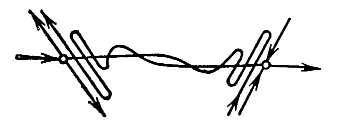


Fig. 101

ment non perturbé et la fréquence résonnante $p/q = 1/3$ tend vers zéro, cette courbe fermée se superpose à la trajectoire fermée initiale en la décrivant trois fois.

La stabilité des positions d'équilibre du système moyennisé admet la même interprétation que celle des solutions périodiques du système perturbé, etc. Il existe toutefois une différence fondamentale dans le cas précisément, où le système moyennisé possède une séparatrice qui joint un col à l'autre.

Dans le système perturbé, aux cols correspond une trajectoire fermée, aux séparatrices rentrante et partante les variétés invariantes attractive et répulsive de cette trajectoire fermée. Mais si dans le système moyennisé les séparatrices se confondent à partir du moment où elles s'intersectent, dans le système perturbé il en va en général autrement. Pour avoir une idée de la façon dont les variétés invariantes se coupent dans l'espace à trois dimensions du système perturbé nous allons étudier la section de cet espace par le plan $t = 0$.

La trajectoire coupe ce plan en trois points qui sont les points fixes du cube de l'application de Poincaré. Chacun de ces points possède une variété invariante rentrante et partante (une courbe). Mais ces courbes invariantes peuvent se couper sans pour autant se confondre (contrairement aux trajectoires de l'équation sur le plan, qui, à partir du moment où elles s'intersectent, sont forcément confondues).

Les images des arcs disjoints des variétés invariantes par les itérations de l'application de Poincaré forment un réseau compliqué appelé figure homoclinique *) (fig. 101).

*) Un point fixe d'un difféomorphisme du plan est homoclinique si les courbes invariantes partante et rentrante se coupent sans se confondre.

I. Résonances d'autres ordres.

Dans le cas des résonances d'ordre $q > 3$, pour système moyennisé de première approximation non triviale on obtient

$$\dot{z} = az + zA(|z|^2) + \bar{z}^{q-1}.$$

En particulier, pour la résonance d'ordre 4, on a

$$\dot{z} = az + Az|z|^2 + \bar{z}^3.$$

Ces systèmes et ceux correspondant à la résonance d'ordre 2 sont examinés en détail dans le chapitre 6.

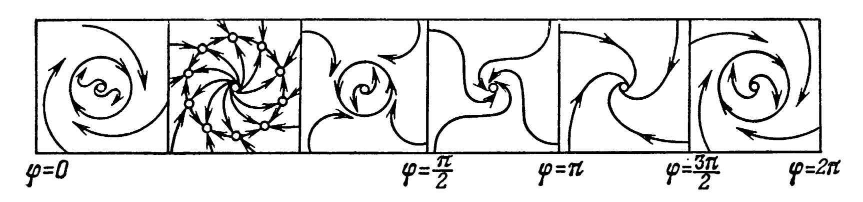


Fig. 102

La figure 102 représente la modification des trajectoires du système moyennisé correspondant à la résonance d'ordre 5

$$\dot{z} = az + Az|z|^2 + \bar{z}^4$$

pour Re $A < 0$, Im $A < 0$, $a = \varepsilon e^{i\varphi}$, $\varepsilon \ll 1$.

CHAPITRE 5

FORMES NORMALES

Il est très fécond non pas de résoudre les équations différentielles, mais de les ramener à une forme plus simple. La théorie des formes normales, due à Poincaré, nous indique les formes les plus simples auxquelles se ramène une équation différentielle au voisinage d'une position d'équilibre ou d'un mouvement périodique.

La réduction aux formes normales s'effectue à l'aide des séries en les écarts par rapport à la position d'équilibre ou au mouvement périodique. Ces séries ne sont pas toujours convergentes. Cependant même dans les cas où les séries sont divergentes, la méthode des formes normales est un puissant outil d'étude des équations différentielles: les premiers termes de la série fournissent sur le comportement des solutions une information suffisante pour leur étude. La méthode des formes normales est aussi le principal outil d'étude de la théorie des bifurcations, où elle s'applique aux familles d'équations dépendant des paramètres.

Dans ce paragraphe on expose les notions fondamentales élémentaires de la méthode des formes normales.

§ 22. Réduction formelle à une forme normale linéaire

Le théorème de Poincaré affirme que dans la classe des séries entières formelles un champ de vecteurs « non résonnant » peut être ramené à sa partie linéaire en un point singulier par un difféomorphisme formel. Enonçons cette condition de non-résonance.

A. Résonances.

Au lieu d'un champ de vecteurs nous allons considérer une série entière vectorielle $v(x) = Ax + \ldots$ de n variables, à coefficients complexes. On suppose que les valeurs propres de la matrice A sont distinctes.

D é f i n i t i o n. Une collection de valeurs propres $\lambda = (\lambda_1, \ldots, \lambda_n)$ est *résonnante* si ces valeurs sont reliées par une relation de

la forme

$$\lambda_s = (m, \lambda),$$

où $m = (m_1, \ldots, m_n)$, $m_k \geqslant 0$, $\sum m_k \geqslant 2$. Cette relation est appelée *résonance*, le nombre $|m| = \sum m_k$, *ordre* de la résonance.

Exemple. La relation $\lambda_1 = 2\lambda_2$ est une résonance d'ordre 2, $2\lambda_1 = 3\lambda_2$ n'est pas une résonance, $\lambda_1 + \lambda_2 = 0$ est une resonance d'ordre 3 (plus exactement, cette relation entraîne la résonance $\lambda_1 = 2\lambda_1 + \lambda_2$).

B. Théorème de Poincaré.

Le théorème suivant constitue le principal résultat de la thèse de Poincaré.

Théorème. *Si les valeurs propres de la matrice A ne sont pas résonnantes l'équation*

$$\dot{x} = Ax + \ldots$$

se ramène par un changement formel des variables $x = y + \ldots$ *à l'équation linéaire*

$$\dot{y} = Ay$$

(*les points de suspension représentent les séries dont les termes sont d'ordre* > 1).

La démonstration du théorème de Poincaré consiste en une élimination successive des termes au carré, au cube, etc. du second membre. Chaque pas est basé sur la résolution d'une équation homologique linéaire dont nous allons donner la déduction.

C. Déduction de l'équation homologique.

Soit h un polynôme vectoriel *) en y d'ordre $r \geqslant 2$ et $h(0) = h'(0) = 0$.

Lemme. *Le changement de variables* $x = y + h(y)$ *transforme l'équation différentielle* $\dot{y} = Ay$ *en l'équation*

$$\dot{x} = Ax + v(x) + \ldots,$$

où $v(x) = \frac{\partial h}{\partial x} Ax - Ah(x)$, *les points de suspension désignent les termes d'ordre* $> r$.

$$◀\ \dot{x} = (E + \partial h/\partial y)\, Ay = (E + \partial h/\partial y)\, A\,(x - h(x) + \quad) =$$
$$= Ax + \left[\frac{\partial h}{\partial x} Ax - Ah(x)\right] + \ldots\ ▶$$

---

*) C'est-à-dire un champ de vecteurs dont les composantes sont des polynômes. Un polynôme vectoriel (un polynôme) est la somme de *monômes vectoriels*; ces derniers étant des champs dont une composante est un monôme, et les autres, des zéros. L'*ordre* du polynôme est la puissance du terme inférieur.

R e m a r q u e. Au second membre on reconnaît le crochet de Poisson des champs de vecteurs Ax et $h(x)$.

On notera L_A l'opérateur associant à tout champ le crochet de Poisson du champ linéaire Ax et de ce champ:

$$L_A h = \frac{\partial h}{\partial x} Ax - Ah(x).$$

D é f i n i t i o n. On appelle *équation homologique attachée à l'opérateur linéaire* A l'équation

$$L_A h = v,$$

où h est un champ de vecteurs inconnu, v un champ de vecteurs connu.

D. Résolution de l'équation homologique.

L'opérateur linéaire L_A opère dans l'espace des champs de vecteurs formels. Il laisse invariant chaque espace de polynômes vectoriels homogènes.

Calculons les valeurs et les vecteurs propres de l'opérateur L_A. Soit e_i vecteur propre de l'opérateur A associé à la valeur propre λ_i. On désignera par $(x_1, \ldots, x_n)$ des coordonnées dans la base $(e_1, \ldots, e_n)$. Comme toujours x^m désignera $x_1^{m_1} \ldots x_n^{m_n}$.

L e m m e. *Si l'opérateur A est diagonal, l'opérateur L_A le sera également sur l'espace des polynômes vectoriels homogènes. Les vecteurs propres de l'opérateur L_A sont les monômes vectoriels $x^m e_s$. Les valeurs propres de l'opérateur L_A dépendent linéairement des valeurs propres de l'opérateur A, plus exactement,*

$$L_A x^m e_s = [(m, \lambda) - \lambda_s]\, x^m e_s.$$

◀ Supposons que $h = x^m e_s$. Alors seule la s-ième composante du vecteur $\frac{\partial h}{\partial x} Ax$ est non nulle; elle vaut

$$\frac{\partial x^m}{\partial x} Ax = \sum \frac{m_i}{x_i} x^m \lambda_i x_i = (m, \lambda)\, x^m.$$

Or $Ah(x) = \lambda_s h(x)$. ▶

L'opérateur L_A est inversible si toutes ses valeurs propres sont non nulles.

C o r o l l a i r e. *Si l'ensemble des valeurs propres de l'opérateur A est non résonnant, l'équation homologique $L_A h = v$ est soluble dans la classe des séries entières formelles h pour tout champ de vecteurs formel sans terme libre et sans partie linéaire au point zéro.*

En l'absence de la résonance d'ordre k, l'équation homologique $L_A h = v$ est résoluble pour tout polynôme vectoriel homogène v de degré k dans la classe des polynômes vectoriels homogènes de degré k (ici $k \geqslant 2$).

R e m a r q u e. Si l'opérateur A n'est pas diagonal (possède des blocs de Jordan), l'opérateur L_A ne le sera pas non plus, mais ses valeurs propres, il est aisé de le voir, sont données par la même

formule que dans le cas diagonal. Donc l'opérateur L_A est inversible sur l'espace des polynômes vectoriels homogènes pour les valeurs propres non résonnantes (même multiples). Donc, le corollaire est valable aussi dans le cas de valeurs propres multiples.

E. Démonstration du théorème de Poincaré.

◀ Supposons que l'équation initiale est de la forme $\dot{x} = Ax + v_r(x) + \ldots$, où v_r sont les termes de degré r ($r \geqslant 2$).

Résolvons l'équation homologique $L_A h_r = v_r$ (en utilisant le corollaire du point D). Faisons la substitution $x = y + h_r(y)$.

L'équation initiale devient $\dot{y} = Ay + w_{r+1}(y) + \ldots$ (on utilise le lemme du point C). Nous avons donc éliminé les termes de degré r au second membre de l'équation initiale.

En éliminant successivement les termes de degré 2, 3, . . . nous obtenons une suite de substitutions dont le produit se stabilise dans la classe des séries formelles, c'est-à-dire les termes de tout degré fixe ne varient pas à partir d'un certain pas. La substitution limite transforme l'équation formelle en l'équation $\dot{y} = Ay$. ▶

Remarque 1. Bien que la convergence des séries n'ait pas été démontrée, par une substitution convergente on peut reculer la perturbation aussi loin que l'on veut dans le cas non résonnant: nous avons montré que pour tout N l'équation initiale peut être ramenée à la forme $\dot{y} = Ay + o(|y|^N)$ par un changement de variables (même polynomial).

Remarque 2. Si la perturbation $v = v_r + v_{r+1} + \ldots$ est d'ordre r, alors en résolvant l'équation homologique $L_A h = v$, on obtient par la substitution $x = y + h$ une équation comportant une perturbation d'ordre $2r - 1$, circonstance qui est liée à la superconvergence des approximations obtenues en itérant cette procédure (cf. § 12).

Remarque 3. La démonstration du théorème de Poincaré est valable pour le cas de valeurs propres multiples (cf. remarque de la fin du point D), pourvu qu'elles ne soient pas résonnantes.

Remarque 4. Si l'équation initiale est réelle et qu'il existe des valeurs propres complexes, la base propre peut être composée de vecteurs conjugués complexes. Dans ce cas toutes les substitutions opérées dans le théorème de Poincaré peuvent être réelles, c'est-à-dire associer des vecteurs conjugués complexes à des vecteurs conjugués complexes.

§ 23. Cas résonnant

Dans le cas résonnant le théorème de Poincaré affirme que l'on peut éliminer tous les termes non résonnants de l'équation par une substitution formelle des variables.

A. Monômes résonnants.

Supposons que l'ensemble des valeurs propres $\lambda = (\lambda_1, \ldots, \lambda_n)$ d'un opérateur A est résonnant. Soient e_s un vecteur de la base propre, x_i les coordonnées dans la base e_i, $x^m = x_1^{m_1} \ldots x_n^{m_n}$, un monôme en x_i.

D é f i n i t i o n. Un monôme vectoriel $x^m e_s$ est résonnant si $\lambda_s = (m, \lambda)$, $|m| \geqslant 2$.

E x e m p l e. L'unique monôme résonnant correspondant à la résonance $\lambda_1 = 2\lambda_2$ est le monôme $x_2^2 e_1$. Tous les monômes $(x_1 x_2)^k x_s e_s$ sont résonnants pour la résonance $\lambda_1 + \lambda_2 = 0$.

B. Théorème de Poincaré-Dulac.

Soit l'équation différentielle

$$\dot{x} = Ax + \ldots$$

définie par la série formelle $v(x) = Ax + \ldots$.

T h é o r è m e. *Cette équation se ramène à la forme canonique*

$$\dot{y} = Ay + w(y),$$

où tous les monômes de la série w sont résonnants, par la substitution formelle $x = y + \ldots$

◀ Eliminons les termes non linéaires de la série v. Au bout de quelques pas on risque de se trouver en présence de l'équation homologique insoluble

$$L_A h = v$$

en le polynôme vectoriel homogène h de degré r égal à l'ordre de la résonance. Dans ce cas il est impossible d'éliminer tous les termes de degré r de la perturbation v par une substitution convenable. Nous allons essayer d'éliminer ceux qui peuvent l'être. En d'autres termes, nous allons écrire v et h sous la forme d'une somme de monômes vectoriels

$$v = \sum v_{m,s} x^m e_s, \quad h = \sum h_{m,s} x^m e_s$$

et poser

$$h_{m,s} = \frac{v_{m,s}}{(m, \lambda) - \lambda_s}$$

pour les m et s, pour lesquels le dénominateur est différent de zéro. Ce faisant, nous définissons le champ h.

Effectuons la substitution ordinaire $x = y + h(y)$ de la démonstration du théorème de Poincaré. Ceci éliminera de l'équation initiale tous les termes de degré r à l'exception des résonnants qui se conservent. L'équation devient

$$\dot{y} = Ay + w_r(y) + \ldots,$$

où w_r n'est composé que de termes résonnants.

La marche à suivre est la suivante. Les termes résonnants de w_r n'influent pas sur l'équation homologique que nous résolvons et sont invariants par les substitutions ultérieures. En effet, l'équation

$$\dot{y}=Ay+w_2(y)+\ldots+w_s(y)+\ldots$$

se transforme en l'équation

$$\dot{z}=Az+w_2(z)+\ldots+w_{s-1}(z)+[w_s(z)-(L_A g_s)(z)]+\ldots$$

par la substitution $y=z+g_s(z)$; le crochet de Poisson de w_2 et de g_s est de degré $s+1$ déjà.

Donc, tous les termes non résonnants de degré s disparaissent par un choix convenable de g_s et la démonstration s'achève comme pour le cas non résonnant. ▶

C. Exemples.

Pratiquement le théorème de Poincaré-Dulac est utilisé généralement pour repérer les termes résonnants d'ordre peu élevé et déplacer la perturbation à des termes d'un certain ordre fini, c'est-à-dire ramener l'équation à la forme

$$\dot{x}=Ax+w(x)+o(|x|^N),$$

(où w est un polynôme composé de monômes résonnants) par un changement de variables non plus formel mais véritable (au besoin polynomial).

Exemple 1. Soit un champ de vecteurs sur un plan, présentant un point singulier du type nœud de résonance $\lambda_1=2\lambda_2$. Le théorème de Poincaré-Dulac permet (formellement) de ramener l'équation à la forme normale

$$\begin{cases}\dot{x}_1=\lambda_1x_1+cx_2^2,\\ \dot{x}_2=\lambda_2x_2.\end{cases}$$

Dans ce cas la forme normale est polynomiale car les termes résonnants sont en nombre fini (1 en tout).

Exemple 2. Soit un champ de vecteurs sur le plan $\mathbb{R}^2$, présentant un point singulier de valeurs propres imaginaires purs $\lambda_{1,2}=\pm i\omega$ (un centre en approximation linéaire).

Passons à la base propre. Les vecteurs propres peuvent être pris conjugués complexes. Généralement sur $\mathbb{C}^2$ les coordonnées sont désignées par z et $\bar{z}$ dans cette base (ces nombres ne sont effectivement conjugués que sur le plan réel $\mathbb{R}^2\subset\mathbb{C}^2$).

L'équation différentielle envisagée sur $\mathbb{R}^2$ définit sur $\mathbb{C}^2$ une équation qui s'écrit

$$\dot{z}=\lambda z+\ldots,\qquad \dot{\bar{z}}=\bar{\lambda}\bar{z}+\overline{\ldots}$$

(les points de suspension représentent une série entière en z et $\bar{z}$). On peut ne pas écrire la deuxième équation puisqu'elle est conjuguée de la première.

On a la résonance $\lambda_1+\lambda_2=0$. D'après le théorème de Poincaré-Dulac, notre équation se ramène à la forme

$$\dot{\zeta}=\lambda\zeta+c\zeta|\zeta|^2+O(|\zeta|^5)$$

par un changement *réel* de la variable (cf. remarque 4 du point E, § 23). Donc $r^2=|\zeta|^2$ est une fonction différentiable réelle sur $\mathbb{R}^2$. On a pour cette fonction

$$(r^2)^{\cdot}=\dot{\zeta}\bar{\zeta}+\zeta\dot{\bar{\zeta}}=(2\,\mathrm{Re}\,c)\,r^4+O(r^6).$$

Si la partie réelle c est négative (resp. positive) la position d'équilibre est stable (resp. instable).

Donc, les premiers pas de la méthode de Poincaré nous fournissent une méthode de résolution du problème de la stabilité d'un point singulier neutre en approximation linéaire. La possibilité de poursuivre cette construction et la convergence de cette procédure importent peu, l'essentiel c'est que le « décrément non linéaire » Re c soit différent de zéro.

Remarque. Le théorème de Poincaré est généralisé en théorie des algèbres de Lie par le théorème de Cartan des répliques qui généralise aussi le théorème de la forme normale de Jordan.

Soit une algèbre de Lie de dimension finie. Soit u un élément de cette algèbre. La commutation avec cet élément définit un opérateur linéaire de l'espace des algèbres de Lie dans lui-même $v \longmapsto [u, v]$. L'opérateur u est *semi-simple* si l'opérateur de commutation avec u est diagonal (c'est-à-dire possède une base propre). L'élément u est nilpotent si l'opérateur de commutation avec u est *nilpotent* (c'est-à-dire toutes les valeurs propres de cet opérateur sont nulles).

Le théorème des répliques affirme que tout élément de l'algèbre se décompose (de façon unique) en une somme d'un élément semi-simple S et d'un élément nilpotent N commutant avec lui:

$$u = S + N, \quad SN = NS.$$

Les éléments S et N sont dits *répliques* de l'élément u.

[En théorie de la forme normale de Jordan, S est un opérateur de matrice diagonale, N la somme de blocs de Jordan nilpotents.]

En algèbre de Lie des jets de champs de vecteurs s'annulant en l'origine, les champs semi-simples sont des champs qui dans un système de coordonnées convenable sont linéaires et définis par une matrice diagonale. Un champ nilpotent est constitué d'une partie linéaire nilpotente et de termes de degré supérieur. La condition de commutation de S et N signifie précisément que la partie non linéaire peut être composée uniquement de termes résonnants dans le système de coordonnées mentionné.

Le théorème de Poincaré-Dulac pourrait être déduit du théorème général des répliques (qu'il faut appliquer aux algèbres de Lie des jets de champs de vecteurs au point zéro).

§ 24. Domaine de Poincaré et de Siegel

Lorsqu'on étudie la convergence des séries de Poincaré construites dans les paragraphes précédents, on distingue deux cas selon la disposition des valeurs propres sur le plan de la variable complexe.

A. Plans résonnants.

Soit l'espace complexe à n dimensions de toutes les collections de valeurs propres: $\mathbb{C}^n = \{\lambda = (\lambda_1, \ldots, \lambda_n)\}$.

Définition. On appelle *plan résonnant* l'hyperplan de $\mathbb{C}^n$ défini par l'équation à coefficients entiers

$$\lambda_s = (m, \lambda) \quad m_k \geqslant 0, \quad \sum m_k \geqslant 2.$$

En faisant varier le vecteur entier m et l'indice s, on obtient un nombre dénombrable de plans résonnants. Voyons comment est disposé l'ensemble des plans résonnants dans l'espace des valeurs propres $\mathbb{C}^n$. On trouve que cet ensemble est discret dans une partie de $\mathbb{C}^n$ et partout dense dans l'autre.

D é f i n i t i o n. Une collection de valeurs propres λ appartient au *domaine de Poincaré* si l'enveloppe convexe des n points $(\lambda_1, \ldots, \lambda_n)$ du plan de la variable complexe ne contient pas le zéro.

Une collection de valeurs propres λ appartient au domaine de Siegel si le zéro est contenu dans l'enveloppe convexe des n points $(\lambda_1, \ldots, \lambda_n)$.

R e m a r q u e. Pour $n > 2$ les domaines de Poincaré et de Siegel contiennent des ouverts et leur frontière est un cône appartenant au domaine de Siegel. Pour $n = 2$ le domaine de Siegel est de codimension réelle 1 dans $\mathbb{C}^2$.

B. Résonance dans le domaine de Poincaré.

Supposons qu'une collection de valeurs propres λ appartient au domaine de Poincaré.

T h é o r è m e 1. *Tout point du domaine de Poincaré vérifie un nombre au plus dénombrable de relations de résonance* $\lambda_s = (m, \lambda)$, $|m| \geqslant 2$, $m_i \geqslant 0$ *et possède un voisinage n'intersectant pas les autres plans résonnants.*

En d'autres termes, les plans résonnants sont discrets dans le domaine de Poincaré.

◀ Par définition, il existe dans le plan des nombres complexes une droite réelle qui sépare la collection des valeurs propres du zéro. Considérons les projections orthogonales des valeurs propres sur la normale à cette droite issue du point zéro. Ces projections sont toutes non inférieures à la distance de la droite séparatrice au zéro.

Or les coefficients m_i de la relation de résonance sont non négatifs. Donc pour $|m|$ assez grand la projection de (m, λ) sur la normale sera supérieure à la plus grande projection de la valeur propre sur la normale à la droite séparatrice. ▶

T h é o r è m e 2. *Si les valeurs propres λ de la partie linéaire d'un champ v en O sont contenues dans le domaine de Poincaré, alors même dans le cas résonnant le champ se ramène à une forme normale polynômiale par un changement de variables formel.*

◀ En vertu du théorème 1, les termes résonnants sont en nombre fini, donc le théorème 2 découle du théorème 1 et du théorème de Poincaré-Dulac. ▶

R e m a r q u e. Dans le domaine de Poincaré la résonance n'est possible que dans le cas où l'une des valeurs propres à coefficients non négatifs s'exprime en fonction des autres sans compter cette valeur, c'est-à-dire si $\lambda_s = (m, \lambda)$, alors $m_s = 0$. En effet, si $m_s > 0$, alors $0 = (m, \lambda) - \lambda_s$ possède une projection positive sur la normale à la droite séparatrice.

C. Résonances dans le domaine de Siegel.

Supposons maintenant que la collection de valeurs propres appartient au domaine de Siegel.

T h é o r è m e 3. *Les plans résonnants sont partout denses dans le domaine de Siegel.*

◀ Le point 0 est contenu soit dans un triangle de sommets (λ_1, λ_2, λ_3) soit dans le segment (λ_1, λ_2). Considérons dans le premier cas l'angle de sommet 0 constitué par les combinaisons linéaires de λ_1 et λ_2 avec des coefficients réels non négatifs.

Les multiples négatifs de λ_3 sont situés dans cet angle. Partageons cet angle en parallélogrammes de sommets aux combinaisons linéaires entières des nombres λ_1 et λ_2. Soit d le diamètre d'un tel parallélogramme. Le nombre $N\lambda_3$ est contenu dans un de ces parallélogrammes pour tout N naturel. Il se trouve donc à une distance $\leqslant d$ de l'un des sommets, de sorte que

$$| N\lambda_3 + m_1\lambda_1 + m_2\lambda_2 | \leqslant d.$$

Cette inégalité entraîne que la distance du point λ à la surface résonnante $\lambda_3 = m_1\lambda_1 + m_2\lambda_2 + (N + 1)\,\lambda_3$ est $\leqslant d/N$. Donc, le théorème est prouvé si 0 est contenu dans le triangle.

Si $0 \in \,]\lambda_1, \lambda_2[$, il existe des entiers p_1 et p_2 aussi grands que l'on veut tels que $| p_1\lambda_1 + p_2\lambda_2 | \leqslant d$.

Ceci nous donne une surface résonnante à une distance $d/| p |$ de λ. ▶

D é f i n i t i o n. *Un point* $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^n$ *est de type* (C, ν) si pour tout s

$$| \lambda_s - (m, \lambda) | \geqslant C/| m |^\nu$$

pour tous les vecteurs entiers m à composantes entières non négatives m_i, $\sum m_i = | m | \geqslant 2$.

T h é o r è m e 4. *L'ensemble des points qui pour aucun C ne sont du type (C, ν) est de mesure nulle si $\nu > (n - 2)/2$.*

◀ Fixons une boule dans $\mathbb{C}^n$ et estimons la mesure de ses points qui ne sont pas de type (C, ν). L'inégalité de la définition définit un voisinage de la surface résonnante d'une largeur $\leqslant C_1C/| m |^{\nu+1}$. Donc la mesure de l'intersection de ce voisinage et de la boule est $\leqslant C_2C^2/ | m |^{2\nu+2}$. En sommant sur m avec $| m |$ fixe, on obtient $| m |^{n-1}\, C_3C^2/| m |^{2\nu+2}$. En sommant sur $| m |$, il vient $C_4(\nu)\, C^2 < \infty$ si $\nu > (n - 2)/2$. Donc, l'ensemble des points qui ne sont pas de type (C, ν) est recouvert par des ensembles de mesure aussi petite que l'on veut. ▶

Dans le cas réel, on exige que $\nu > n - 1$ dans le théorème 4.

D. Théorèmes de Poincaré et de Siegel.

Supposons maintenant que le champ de vecteurs est défini par une série convergente, c'est-à-dire nous considérons une équation différentielle à second membre holomorphe.

T h é o r è m e d e P o i n c a r é. *Si les valeurs propres de la partie linéaire d'un champ de vecteurs holomorphe en un point singulier*

appartiennent au domaine de Poincaré et ne sont pas résonnants, alors ce champ est biholomorphement équivalent à sa partie linéaire au voisinage du point singulier.

Il est équivalent de dire que les séries de Poincaré construites dans les paragraphes précédents sont convergentes si les valeurs propres appartiennent au domaine de Poincaré.

Théorème de Siegel. *Si les valeurs propres de la partie linéaire d'un champ de vecteurs holomorphe sont les composantes d'un vecteur de type (C, ν) en un point singulier, alors ce champ est biholomorphement équivalent à sa partie linéaire au voisinage de ce point singulier.*

Autrement dit, les séries de Poincaré sont convergentes pour presque toutes (au sens de la théorie des mesures) les parties linéaires du champ au point singulier.

Remarque. Tous les vecteurs non résonnants du domaine de Poincaré sont des vecteurs de type (C, ν) avec $C > 0$. Au contraire, sont partout denses dans le domaine de Siegel aussi bien les (C, ν)-vecteurs et les vecteurs résonnants que les vecteurs non résonnants qui ne sont du type (C, ν) pour aucuns C et ν.

Pour des valeurs propres du dernier type, (à composantes incommensurables mais assez proches de la commensurabilité), les séries de Poincaré peuvent diverger de sorte que le champ peut être formellement équivalent à sa partie linéaire, mais biholomorphement non équivalent.

Les démonstrations des théorèmes de Poincaré et de Siegel se déduisent moyennant quelques simplifications de celles de théorèmes analogues pour les applications dont on trouvera les démonstrations au § 28.

E. Théorème de Poincaré-Dulac.

Etudions maintenant le cas des valeurs propres résonnantes.

Théorème. *Si les valeurs propres de la partie linéaire d'un champ de vecteurs holomorphe en un point singulier appartiennent au domaine de Poincaré, alors le champ est, au voisinage de ce point singulier, biholomorphement équivalent à un champ de vecteurs polynômial dont tous les monômes à coefficients supérieurs à un sont résonnants.*

En d'autres termes, les séries de Poincaré sont convergentes si les valeurs propres sont situées dans le domaine de Poincaré même en cas de résonance.

Remarque. Si au contraire les valeurs propres sont situées dans le domaine de Siegel, les séries qui conduisent aux formes normales formelles en cas de résonance sont souvent divergentes. Le premier exemple du genre a été construit par Euler (L. Euler, De seriebus divergenti bus, Opera omnia, Ser. 1, 14 (1924), Leipzig-Berlin, 247, 585-617; voir page 601).

Dans l'exemple d'Euler

$$\begin{cases} \dot{x} = x^2, \\ \dot{y} = y - x \end{cases}$$

l'origine des coordonnées est un point singulier du type col-nœud. Malgré l'analyticité du second membre, la séparatrice du demi-plan $x < 0$ n'est pas analytique mais seulement indéfiniment différentiable: $y = \sum (k-1)!\, x^k$.

A. Bruno a exhibé de nombreux exemples de séries divergentes de Poincaré (A. Bruno, Forme analytique des équations différentielles, Troudy MMO 25 (1971), 119-262; dans ce travail est également démontrée la convergence des séries dans quelques cas n'entrant pas dans le cadre du théorème de Siegel).

F. Cas réel et non analytique.

Les théorèmes de Poincaré et Poincaré-Dulac se généralisent au cas réel et analytique, au cas de champs de vecteurs indéfiniment différentiables et même au cas de champs différentiables un nombre de fois assez grand.

Une telle généralisation est également possible au cas de Siegel (i.e. les valeurs propres appartiennent au domaine de Siegel) (cf., par exemple, S. Sternberg, On the structure of local homeomorphisms of Euclidean n-space, Amer. J. Math. 80, 3 (1958), 623-631, 81, 3 (1959), 578-604).

A noter toutefois que les cas justiciables de ces théorèmes sont topologiquement triviaux. En effet, le cas de Poincaré (i.e. les valeurs propres appartiennent au domaine de Poincaré) pour un champ réel se présente lorsque les valeurs propres sont toutes situées soit dans le demi-plan de droite, soit dans le demi-plan de gauche. Dans ce cas (indépendamment des résonances) le système est topologiquement équivalent au système standard $\dot{x} = -x$ (ou $\dot{x} = +x$) au voisinage d'un point fixe de l'espace réel. Toutes les trajectoires tendent vers une position asymptotiquement stable pour $t \to +\infty$ (ou bien ont quitté la position d'équilibre pour $t \to -\infty$).

Dans la situation du théorème de Siegel on peut se servir du théorème de Grobman-Hartman (le système est topologiquement équivalent à un col standard) dans le domaine réel. En effet, si l'une au moins des valeurs propres non nulles de la partie linéaire est située sur l'axe imaginaire, la valeur propre conjuguée complexe est également située sur l'axe imaginaire; le couple $\lambda_{1,2} = \pm i\omega$ conduit à la résonance $\lambda_1 + \lambda_2 = 0$. La valeur propre zéro est toujours résonnante. Donc, le théorème de Siegel s'applique dans le cas réel uniquement aux systèmes sans valeurs propres sur l'axe imaginaire, or ces systèmes sont localement topologiquement équivalents à leur partie linéaire (théorème de Grobman-Hartman, § 13).

A l'inverse des *théorèmes* de Poincaré et de Siegel, la méthode de Poincaré s'applique à l'étude des cas topologiquement compliqués où les valeurs propres sont situées sur l'axe imaginaire. Plus exactement, cette méthode est utilisée pour la normalisation d'un nombre fini de termes de la série de Taylor. On démontre ensuite que les termes d'ordre supérieur ne changent rien à la situation.

Un exemple simple de cette nature est examiné plus haut au point C, § 23. Cette méthode est très utile en théorie des bifurcations (cf. chapitre 6).

§ 25. Forme normale d'une application au voisinage d'un point fixe

La construction d'un système de coordonnées convenable pour l'application d'un espace dans lui-même au voisinage d'un point fixe rappelle la théorie des formes normales des équations différentielles au voisinage de la position d'équilibre. Dans ce paragraphe on indique comment se formulent en conséquence les notions fondamentales de la théorie des formes normales.

A. Résonances. Domaines de Poincaré et de Siegel.
Soient une application formelle $F: \mathbb{C}^n \to \mathbb{C}^n$ définie par la série entière $F(x) = Ax + \ldots$, $(\lambda_1, \ldots, \lambda_n)$ les valeurs propres de l'opérateur linéaire A. On appelle *résonance* la relation

$$\lambda_s = \lambda^m, \text{ où } \lambda^m = \lambda_1^{m_1} \ldots \lambda_n^{m_n}, \; m_k \geqslant 0, \; \sum m_k \geqslant 2.$$

Exemple. Pour $n = 1$ sont résonnantes la valeur 0 et les racines de toute puissance entière de l'unité, sont non résonnantes les autres valeurs λ.

Définition. Une collection de valeurs propres appartient au domaine de Poincaré si leurs modules sont tous plus petits ou tous plus grands que l'unité.

Donc, si une application F est telle que les valeurs propres de sa partie linéaire appartiennent au domaine de Poincaré, alors au voisinage de l'origine des coordonnées ou bien elle, ou bien son inverse est contractante selon que $|\lambda|$ est < 1 ou > 1.

Définition. Le complément du domaine de Poincaré est le domaine de Siegel. Pour $n = 1$ le domaine de Siegel est le cercle unité $|\lambda| = 1$. L'équation de la résonance $\lambda_s = \lambda^m$ définit une hypersurface complexe dans l'espace des valeurs propres $\mathbb{C}^n$, appelée *surface résonnante*. Les surfaces résonnantes sont discrètes dans le domaine de Poincaré. Les points résonnants et les points non résonnants sont partout denses dans le domaine de Siegel.

B. Linéarisation formelle.
On étudie d'abord la forme normale formelle au point fixe d'une application.

Théorème. *Si la collection des valeurs propres d'une application F n'est pas résonnante au point fixe de F, alors l'application $x \mapsto F(x)$ se ramène à sa partie linéaire $x \mapsto Ax$ par le changement formel de variables $x = \mathcal{H}(y) = y + \ldots$:*

$$F \circ \mathcal{H} = \mathcal{H} \circ A.$$

◀ Soit $H(y) = y + h(y)$, où h est un polynôme vectoriel homogène de degré $r \geqslant 2$. On a

$$H \circ A \circ H^{-1}(x) = Ax + [h(Ax) - Ah(x)] + \ldots,$$

où les points de suspension désignent les termes d'ordre plus grand que r. L'expression entre crochets est un polynôme vectoriel homogène de degré r, dépendant linéairement de h. L'opérateur linéaire

$$M_A : h(x) \mapsto [h(Ax) - Ah(x)]$$

possède sur l'espace des polynômes vectoriels homogènes les valeurs propres $\lambda^m - \lambda_s$ et les vecteurs propres $h(x) = x^m e_s$ (comme toujours les vecteurs e_k constituent la base propre de l'opérateur A, $x^m = x_1^{m_1} \ldots x_n^{m_n}$, x_k, les coordonnées dans la base $\{e_k\}$; les valeurs propres de l'opérateur, par souci de simplicité, sont supposées distinctes). Nous sommes donc conduits à l'équation homologique

$$M_A h = v;$$

pour résoudre cette équation il faut diviser les coefficients du développement de v par les nombres $\lambda^m - \lambda_s$. Donc, la condition de résonance s'écrit ici $\lambda_s = \lambda^m$.

La suite de la démonstration du théorème est la même que celle effectuée au § 22 pour les équations différentielles. ▶

C. Problèmes de convergence.

Les théorèmes de Poincaré et de Siegel admettent les généralisations suivantes au cas où le temps est discret.

Théorème de Poincaré. *Si les valeurs propres d'un difféomorphisme holomorphe en un point fixe sont toutes inférieures (ou toutes supérieures) à l'unité en module et si n'existe pas de résonance, alors cette application se transforme en sa partie linéaire par un difféomorphisme biholomorphe local au voisinage du point fixe.*

Théorème de Siegel. *Pour presque toute (au sens de la mesure de Lebesgue) collection des valeurs propres de la partie linéaire d'un difféomorphisme holomorphe en son point fixe, ce difféomorphisme se transforme en sa partie linéaire par un difféomorphisme biholomorphe en ce point.*

Plus exactement, pour qu'un difféomorphisme soit équivalent à sa partie linéaire, il suffit que ses valeurs propres vérifient les inégalités

$$|\lambda_s - \lambda^m| \geqslant C |m|^{-\nu}$$

pour tous les $s = 1, \ldots, n$, $|m| = \sum m_k \geqslant 2$, $m_k \geqslant 0$. Les collections de valeurs propres vérifiant cette inégalité sont dites collections de type multiplicatif (C, ν). L'ensemble des collections de valeurs propres λ qui ne sont pas de type multiplicatif (C, ν) pour aucun C est de mesure nulle si $\nu > (n - 1)/2$.

Les théorèmes de Poincaré et de Siegel se démontrent pratiquement de la même façon que pour les équations différentielles. Bien que le théorème de Siegel soit déjà connu depuis plus de 30 ans sa démonstration n'a semble-t-il pas été publiée à ce jour. Une démonstration est donnée au § 28.

D. Cas résonnant.

A toute résonance $\lambda_s = \lambda^m$ est associé un monôme vectoriel résonnant $x^m e_s$ (où e_s est un vecteur de la base propre, $x^m = x_1^{m_1} \ldots x_n^{m_n}$, x_k les coordonnées dans la base propre).

Théorème de Poincaré-Dulac. *L'application formelle $x \mapsto Ax + \ldots$, où la matrice de l'opérateur A est diagonale, se ramène par le changement de variables formel $x = y + \ldots$ à la forme normale $y \mapsto Ay + w(y)$, où la série w est constituée uniquement de monômes résonnants. Si les valeurs propres de la partie linéaire de l'opérateur A sont toutes inférieures (ou toutes supérieures) en module à l'unité, alors l'application holomorphe $x \mapsto Ax + \ldots$ se ramène par un changement de variables biholomorphe à une forme normale polynômiale composée uniquement de termes résonnants.*

Dans le cas résonnant la méthode de Poincaré est généralement utilisée pour réduire à la forme normale un nombre fini de termes de la série de Taylor d'une application en son point fixe.

Exemple. Soit une application de $\mathbb{C}^1$ dans $\mathbb{C}^1$, de point fixe O, et dont la valeur propre λ est racine n-ième de l'unité. Cette application s'écrit dans un système de coordonnées convenable

$$x \mapsto \lambda x + cx^{n+1} + O(|x|^{2n+1}).$$

Pour $\lambda = -1$ par exemple, cette application devient

$$x \mapsto -x + cx^3 + O(|x|^5).$$

Cette formule permet d'étudier la stabilité du point fixe d'une application réelle. En effet, le carré de cette application est

$$x \mapsto x - 2cx^3 + O(|x|^5).$$

Donc, si $c > 0$, le point fixe O est stable.

Par suite, les premiers pas de la méthode de Poincaré permettent d'étudier la stabilité du point fixe dans le cas où l'approximation linéaire est douteuse.

§ 26. Forme normale d'une équation à coefficients périodiques

Une variante de la méthode des formes normales de Poincaré permet de réduire une équation à coefficients périodiques à sa forme simple.

A. Forme normale d'une équation linéaire à coefficients périodiques.

Soit dans un espace des phases complexe l'équation linéaire

$$\dot{x} = A(t)\, x,$$

où l'opérateur linéaire complexe $A(t)\colon \mathbb{C}^n \to \mathbb{C}^n$ est 2π-périodique en t.

On appelle *opérateur de monodromie* un opérateur linéaire $M\colon \mathbb{C}^n \to \mathbb{C}^n$ associant à la condition initiale pour $t = 0$ la valeur prise

au point $t = 2\pi$ par la solution vérifiant cette condition initiale (la monodromie est définie non seulement pour les équations linéaires, mais aussi pour toutes les équations à coefficients périodiques; dans ce cas plus général la monodromie est appelée *application de Poincaré*).

Théorème de Floquet. *Si l'opérateur de monodromie est diagonal et $\mu_s = e^{2\pi\lambda_s}$ sont ses valeurs propres, alors l'équation linéaire initiale à coefficients périodiques se ramène à l'équation à coefficients constants*

$$\dot{y} = \Lambda y,$$

où Λ est un opérateur diagonal de valeurs propres λ_s, par un changement de variables linéaire 2π-périodique $x = B(t)\, y$.

◀ Soit un opérateur linéaire associant à la condition initiale de l'équation initiale pour $t = 0$ la valeur prise à l'instant t par la solution vérifiant cette condition initiale. Désignons cet opérateur par g^t: $\mathbb{C}^n \to \mathbb{C}^n$. Soit f^t: $\mathbb{C}^n \to \mathbb{C}^n$ l'opérateur analogue pour l'équation $\dot{y} = \Lambda y$. Alors $g^0 = f^0 = E$, $g^{2\pi} = f^{2\pi} = M$ est opérateur de monodromie (en vertu du choix de Λ). Posons $B(t) = g^t(f^t)^{-1}$. L'opérateur $B(t)$ définit le changement cherché. ▶

Remarque. Pour démontrer le théorème de Floquet on s'est servi seulement de la représentation $M = e^{2\pi\Lambda}$ de l'opérateur de monodromie. Donc un changement de variables périodique ramène à une équation à coefficients constants non seulement une équation complexe à opérateur de monodromie diagonal mais aussi toute équation pour laquelle l'opérateur de monodromie possède un logarithme.

Tout opérateur linéaire complexe non dégénéré possède un logarithme (il est aisé de s'en assurer en écrivant la matrice de l'opérateur sous la forme de Jordan).

Corollaire 1. *Toute équation linéaire complexe à coefficients 2π-périodiques se ramène à une équation à coefficients constants par un changement de variables linéaire 2π-périodique.*

Un opérateur linéaire réel ne possède pas toujours de logarithme réel même si son déterminant est positif (le déterminant d'un opérateur de monodromie est toujours positif). En effet, considérons par exemple un opérateur linéaire sur le plan de valeurs propres (—1, —2). Si cet opérateur est l'exponentielle d'un autre opérateur linéaire, alors les valeurs propres de ce dernier sont complexes mais ne sont pas des nombres conjugués complexes. Donc, notre opérateur ne possède pas de logarithme sur le plan réel.

Par ailleurs, il est aisé de vérifier que le carré d'un opérateur linéaire réel possède toujours un logarithme réel. D'où le

Corollaire 2. *Toute équation linéaire réelle à coefficients 2π-périodiques se ramène à une équation à coefficients constants par un changement de variables linéaire 4π-périodique.*

La réduction complexe est généralement plus commode que la réduction réelle à double période.

B. Déduction de l'équation homologique.

Soit une équation linéaire à coefficients constants $\dot{y} = \Lambda y$. Effectuons un changement de coordonnées non linéaire 2π-périodique en t

$$x = y + h(y, t),$$

où h est une fonction vectorielle (ou une série formelle sur les puissances de y) à coefficients 2π-périodiques.

L e m m e. *Si* $h = O(|y|^r)$ *(ou la série* h *commence par des termes de degré* $\geqslant r$*),* $r \geqslant 2$, *alors*

$$\dot{x} = \Lambda x + \left[\frac{\partial h}{\partial x}\Lambda x - \Lambda h + \frac{\partial h}{\partial t}\right] + \ldots$$

où les points de suspension désignent les termes en x *de degré* $> r$.

◀ $\dot{x} = (E + h_y)\Lambda y + h_t = (E + h_y)\Lambda(x - h(x, t)) + h_t + \ldots =$
$= \Lambda x + [h_x\Lambda x - \Lambda h(x, t) + h_t] + \ldots$ ▶

D é f i n i t i o n. On appelle *équation homologique* rattachée à l'équation à coefficients 2π-périodiques $\dot{y} = \Lambda y$ l'équation en un champ de vecteurs h 2π-périodique en t

$$L_\Lambda h + h_t = v,$$

où v est un champ de vecteurs 2π-périodique

$$(L_\Lambda h)(x, t) = \frac{\partial h}{\partial x}\Lambda x - \Lambda h(x, t).$$

Nous traiterons les cas où h et v sont des séries formelles à coefficients 2π-périodiques en t.

C. Résolution de l'équation homologique.

Supposons d'abord que v et h sont des séries de Taylor-Fourier

$$v(x, t) = \sum v_{m, k, s} x^m e^{ikt} e_s,$$

$$h = \sum h_{m, k, s} x^m e^{ikt} e_s.$$

La solution formelle de cette équation est

$$h_{m, k, s} = \frac{v_{m, k, s}}{ik + (m, \lambda) - \lambda_s},$$

où λ_j sont les valeurs propres de l'opérateur Λ.

La condition de résonance est

$$\lambda_s = (m, \lambda) + ik,$$
$$m_j \geqslant 0, \ \sum m_j \geqslant 2, \quad -\infty < k < +\infty,$$
$$1 \leqslant s \leqslant n.$$

S'il n'y a pas de résonance pour m, s donnés, alors la série de Fourier $\sum h_{m,\,k,\,s} e^{ikt}$ et sa dérivée par rapport à t sont convergentes. Donc, s'il n'y a pas de résonance, les solutions de l'équation homologique appartiennent à la classe des polynômes homogènes à coefficients 2π-périodiques en t et partant à la classe des séries formelles entières à coefficients 2π-périodiques en t.

S'il y a résonance, l'équation homologique est formellement soluble dans le cas où la série de Taylor-Fourier pour v ne contient pas de termes résonnants, c'est-à-dire sont nuls les coefficients $v_{m,\,k,\,s}$ pour tous les termes de la série pour lesquels est réalisée la condition de résonance $\lambda_s = ik + (m, \lambda)$.

D. Forme normale formelle.

Dans le cas non résonnant, on ramène l'équation à coefficients formels 2π-périodiques selon la procédure habituelle à une équation linéaire à coefficients constants $\dot{y} = \Lambda y$ par un changement de variables ayant la forme d'une série formelle sur y à coefficients 2π-périodiques en t.

Dans le cas résonnant nous ramenons l'équation à la forme

$$\dot{y} = \Lambda y + w(y, t),$$

où w est une série formelle sur les puissances de y à coefficients 2π-périodiques en t, composée uniquement de termes résonnants (on remarquera que les termes résonnants de tout ordre fixe sur y ne contiennent qu'un nombre fini d'harmoniques de Fourier, puisque la condition de résonance $\lambda_s = (m, \lambda) + ik$ définit k de façon unique).

En général on n'utilise pratiquement que la normalisation des termes d'ordre peu élevé.

E x e m p l e. Soit une équation à coefficients 2π-périodiques. Supposons que l'espace des phases est de dimension 2 et que les deux valeurs propres de l'opérateur de monodromie sont complexes et égales à l'unité en module.

Dans un système de coordonnées convenable l'équation complexifiée linéarisée s'écrit

$$\dot{z} = i\omega z$$

(on omet d'écrire l'équation en $\bar{z}$ qui est l'équation conjuguée). Les valeurs propres sont $\lambda_{1,\,2} = \pm i\omega$. Les termes résonnants de l'équation en z se déduisent à partir de la condition

$$ik + (m_1 - m_2 - 1)\, i\omega = 0.$$

Si le nombre réel ω est irrationnel, alors $k = 0$, $m_1 = m_2 + 1$. Donc, l'équation se ramène à la forme normale formelle qui ne dépend pas du temps

$$\dot{z} = i\omega z + c_1 z \mid z \mid^2 + c_2 z \mid z \mid^4 + \ldots$$

Par un changement de variables (non formel) on peut ramener l'équation par exemple à la forme

$$\dot{z} = i\omega z + c_1 z \mid z \mid^2 + \ldots,$$

où seuls les termes d'ordre 5 de petitesse en z, désignés par les points de suspension, dépendent 2π-périodiquement de t.

On remarquera que dans ce cas tout pas de la méthode de Poincaré se ramène à une moyennisation sur t et $\arg z$ et que l'équation obtenue est invariante par les translations sur t et les rotations sur z.

E. Cas de commensurabilité.

Supposons maintenant que dans l'exemple précédent le nombre ω est rationnel, c'est-à-dire $\omega = p/q$. Dans ce cas, de l'équation pour les termes résonnants il vient

$$k = pr, \quad m_1 = m_2 + 1 - qr.$$

Pour étudier la forme normale il est commode de considérer le recouvrement à q feuillets de l'axe temporel. On remarquera que les courbes intégrales de la partie linéaire de notre équation constituent un feuilletage de Seifert du type (p, q) (cf. § 21). Les courbes intégrales forment un fibré trivial sur le recouvrement à q feuillets, nous pouvons donc introduire les coordonnées du produit direct. On désignera par t (mod $2\pi q$) la coordonnée sur une fibre. La coordonnée sur la base, ζ, se déduit à partir de la condition

$$z = e^{i\omega t}\zeta.$$

Dans ces notations la partie linéaire de notre équation s'écrit $\dot{\zeta} = 0$, et la forme normale devient une série formelle ne dépendant pas de t

$$\dot{\zeta} = \sum w_{k,\,l}\zeta^k\bar{\zeta}^l,$$

où $k - l \equiv 1 \bmod q$.

En d'autres termes, sur la base du recouvrement à q feuillets on obtient une équation (formelle) invariante par les rotations d'angle $2\pi/q$.

Si, au lieu d'une réduction formelle complète, on se limite à la normalisation de quelques premiers termes de la série, on obtient pour ζ une équation à résidu d'ordre $q + 1$ $2\pi q$-périodique en t:

$$\dot{\zeta} = \zeta a (\mid \zeta \mid^2) + b\bar{\zeta}^{q-1} + \ldots$$

Dans ce cas, tout pas de la méthode de Poincaré se ramène à une moyennisation sur un feuilletage de Seifert, donc l'équation obtenue est invariante par les translations sur t et les rotations sur ζ d'angles $2\pi/q$.

Les équations obtenues sont étudiées au chapitre 6.

F. Discussion de la convergence.

Le domaine de Poincaré pour une équation à coefficients périodiques $\dot{x} = \Lambda x + \ldots$ est défini par la condition : toutes les valeurs propres de l'équation linéarisée sont situées dans le demi-plan de gauche $\operatorname{Re} \lambda < 0$ (ou dans celui de droite).

Dans ce domaine :

1) les plans résonnants $\{\lambda: \lambda_s = (m, \lambda) + ik\}$ sont discrets;

2) la forme normale en la résonance ne contient qu'un nombre fini de termes;

3) les séries de Poincaré sont convergentes.

Le complément du domaine de Poincaré est le domaine de Siegel. Dans le domaine de Siegel :

1) les plans résonnants forment un ensemble partout dense;

2) les formes normales peuvent contenir une infinité de termes;

3) les séries de Poincaré peuvent être divergentes.

Cependant, *pour presque toutes* (*au sens de la mesure de Lebesgue*) *les collections de valeurs propres* λ *de l'opérateur* Λ, *l'équation différentielle holomorphe* 2π*-périodique en* t $\dot{x} = \Lambda x + \ldots$ *se ramène au voisinage de la solution triviale à la forme normale autonome* $\dot{x} = \Lambda x$ *par une application biholomorphe* 2π*-périodique en* t (*théorème de Siegel pour le cas de coefficients périodiques*).

◀ Démonstration habituelle, cf. § 28. ▶

G. Voisinage d'une trajectoire fermée.

Soit l'équation différentielle autonome $\dot{x} = v(x)$ possédant une solution périodique et, par conséquent, une trajectoire fermée. Tout ce qui a été dit plus haut sur le voisinage de la solution nulle d'une équation à coefficients périodiques est valable pour ce cas.

En effet, au voisinage de la trajectoire fermée on peut choisir les coordonnées de telle sorte que le champ de directions définissant le champ de vecteurs v soit champ de directions d'une équation à coefficients périodiques et, de plus, la dimension de l'espace des phases diminue d'une unité (la coordonnée qui varie le long de la trajectoire sera appelée temps).

R e m a r q u e. Si l'espace des phases est une variété, le voisinage de la trajectoire fermée est susceptible de ne pas être difféomorphe au produit direct d'un cercle par un disque transversal.

E x e m p l e. L'espace des phases est le ruban de Möbius, la trajectoire fermée, son cercle axial.

En général, le voisinage d'un cercle dans une variété ne sera pas produit direct si et seulement si la variété n'est pas orientée et le cercle est son chemin désorientant. Dans ce cas pour passer à une équation à coefficients périodiques on doit recourir à un recouvrement à deux feuillets du cercle initial.

H. Lien avec les applications de Poincaré.

La théorie des formes normales des équations à coefficients périodiques pourrait être déduite de la théorie des formes normales de leurs applications de Poincaré, c'est-à-dire des formes normales des difféomorphismes au voisinage de leurs points fixes. Inversement, l'étude d'un difféomorphisme au voisinage de son point fixe pourrait se ramener à celle d'une équation à coefficients périodiques pour laquelle ce difféomorphisme est une application de Poincaré.

Dans le cas d'une différentiabilité réelle finie et même infinie la construction d'une équation différentielle à coefficients périodiques *) correspondant à une application de Poincaré donnée ne présente pas de grandes difficultés. Dans le cas analytique ou holomorphe la situation est plus compliquée. Cette question équivaut à celle de la trivialité analytique (holomorphe) des fibrés analytiques (holomorphes) au-dessus d'un anneau circulaire sous l'hypothèse de la trivialité topologique. La théorie des faisceaux et des variétés de Stein fournit une réponse à cette question mais à ma connaissance la démonstration n'a encore pas été publiée (je remercie V. Palamodov et Y. Iliachenko pour les éclaircissements apportés à ce sujet). Nous ne nous étendrons pas sur cette théorie d'autant plus que tous les résultats indispensables à l'étude des équations différentielles et des difféomorphismes peuvent ne pas être déduits les uns des autres, mais obtenus indépendamment à l'aide de la même méthode de démonstration.

I. Cas de coefficients quasi périodiques.

La méthode de Poincaré admet une généralisation immédiate au cas de coefficients quasi périodiques. Il s'agit de l'équation

$$\dot{x} = \Lambda x + v(x, \varphi),$$
$$\dot{\varphi} = \omega,$$

où φ est un point d'un tore à r dimensions, ω un vecteur constant, $\Lambda: \mathbb{C}^n \to \mathbb{C}^n$ un opérateur linéaire (ne dépendant pas de φ), v un champ de vecteurs dont la partie linéaire est nulle au point $x = 0$.

On impose aux composantes du vecteur des fréquences ω les conditions ordinaires d'incommensurabilité normale. Les conditions de résonance s'écrivent ici

$$\lambda_s = i(k, \omega) + (m, \lambda),$$

où k parcourt le réseau des points entiers d'un espace à r dimensions, et m vérifie les conditions habituelles $m_p \geqslant 0$, $\sum m_p \geqslant 2$.

La fonction v est supposée analytique (holomorphe) en x et φ, 2π-périodique en $\varphi = (\varphi_1, \ldots, \varphi_r)$. On montre que le système se ramène à la forme

$$\dot{y} = \Lambda y, \qquad \dot{\varphi} = \omega,$$

par une substitution analytique (holomorphe) $x = y + h(y, \varphi)$, 2π-périodique en φ (E. Belaga, Sur la réductibilité d'un système d'équations différentielles ordinaires au voisinage d'un mouvement quasi périodique, DAN SSSR, **143**, 2 (1962) 255-258 (en russe).)

Le point faible de cette théorie est la théorie des équations linéaires à coefficients quasi périodiques: alors que pour les équations à coefficients périodiques on pourrait rendre constante la partie linéaire par une transformation linéaire périodique convenable des

*) En général, cette application ne fait pas partie du flot de l'équation autonome (exemple: un difféomorphisme du cercle de nombre de rotation irrationnel ne se ramène pas à une rotation par un difféomorphisme, cf. § 11).

coordonnées, pour les équations à coefficients quasi périodiques l'indépendance de Λ par rapport à φ est une restriction importante.

J. Réductibilité des équations linéaires à coefficients quasi périodiques.

Par équation linéaire à coefficients quasi périodiques on entend le système

$$\dot{x} = A(\varphi)\,x, \qquad \dot{\varphi} = \omega,$$

où $x \in \mathbb{C}^n$, $\varphi \in T^r$, ω est un vecteur à composantes entières indépendantes, $A(\varphi)$ un opérateur linéaire dans $\mathbb{C}^n$.

Donc, l'équation est définie par le couple (A, ω), où A est une fonction différentiable sur le tore à valeurs opératorielles (ou matricielles si l'on veut), ω un vecteur sur le tore.

Définition. Une équation linéaire à coefficients quasi périodiques est *réductible* si existe une fonction opératorielle (différentiable) sur le tore B telle que la substitution $x = B(\varphi)\,y$ transforme cette équation en une équation à coefficients constants $\dot{y} = Cy$.

Le problème de réductibilité revient donc à établir *si est réductible l'équation linéaire générique.*

On ne sait même pas *si existent dans l'espace fonctionnel des couples analytiques* (A, ω) *des domaines ne contenant pas de systèmes réductibles.*

Le problème de la réductibilité d'équations linéaires (et non linéaires) à coefficients quasi périodiques se pose naturellement lorsqu'on étudie le voisinage d'un tore invariant d'une équation autonome possédant une solution quasi périodique. Généralement, on cherche ce tore à l'aide d'approximations successives qui, dans le cas générique, peuvent en principe être modifiées de manière à obtenir simultanément le tore invariant et à ramener à la forme normale une équation aux variations sur ce tore en évitant ainsi le problème irrésolu de la réductibilité (en somme on utilise la réductibilité pour un problème « non perturbé »).

§ 27. Forme normale au voisinage d'une courbe elliptique

La théorie de Poincaré des formes normales des équations différentielles au voisinage d'un point singulier est similaire à la théorie des formes normales des voisinages des courbes elliptiques sur des surfaces complexes. Dans ce paragraphe on se penche brièvement sur cette théorie qui est une application des méthodes de la théorie des équations différentielles à la géométrie analytique et qui possède elle-même des applications en théorie des équations différentielles (§ 36).

A. Courbes elliptiques.

On appelle *courbe elliptique* une variété complexe à une dimension homéomorphe au tore.

Exemple. Soient le plan de la variable complexe $\mathbb{C}$ et deux nombres complexes (ω_1, ω_2) dont le quotient n'est pas réel. Identifions tout point φ de $\mathbb{C}$ à son image par une ω_1-translation et une ω_2-translation (donc, à tous les points $\varphi + k_1\omega_1 + k_2\omega_2$, où k_1 et k_2 sont des entiers). Le plan $\mathbb{C}$ se transforme par cette identification en la courbe elliptique

$$\Gamma = \mathbb{C}/\omega_1 \mathbb{Z} + \omega_2\mathbb{Z}.$$

Donc, on peut se représenter la courbe elliptique Γ comme un parallélogramme de côtés (ω_1, ω_2) dont on a identifié les points correspondants sur les côtés opposés.

On montre que toutes les courbes elliptiques s'obtiennent par la construction précédente à une équivalence biholomorphe près (ce fait est loin d'être un théorème évident).

Considérons, par exemple, la bande $0 \leqslant \operatorname{Im} \varphi \leqslant \tau$ et collons tous les points φ, $\varphi + 2\pi$ ainsi que les points des bords de la bande en identifiant le point φ au point $\varphi + i\tau + \sigma + 0{,}5 \sin \varphi$ pour φ réels. La variété obtenue s'applique biholomorphement sur la variété quotient $\mathbb{C}/\omega_1\mathbb{Z} + \omega_2\mathbb{Z}$, mais la démonstration n'est pas facile. Pour $\tau \to 0$ le quotient ω_1/ω_2 tend probablement vers le nombre de rotation sous les conditions diophantiennes habituelles.

Les nombres ω_1 et ω_2 s'appellent périodes de la courbe. Si l'on multiplie ces deux périodes par un même nombre complexe on obtient de nouvelles périodes qui définissent une courbe elliptique biholomorphement équivalente à la courbe initiale. Donc on peut toujours choisir les périodes de telle sorte que $\omega_1 = 2\pi$.

Dans ce cas on désignera la deuxième période par ω. On peut toujours admettre que $\operatorname{Im} \omega > 0$. A différents ω correspondent généralement des courbes elliptiques biholomorphement non équivalentes (plus exactement, des courbes sont biholomorphement non équivalentes si les réseaux correspondants $\omega_1\mathbb{Z} + \omega_2\mathbb{Z}$ ne se transforment pas l'un dans l'autre par une multiplication par un nombre complexe).

Exercice *). Prouver que les trajectoires de l'équation de Newton à une dimension, d'énergie potentielle de degré 3 ou 4, sont des courbes elliptiques (si on se place dans le domaine complexe).

Indication. La coordonnée φ sur le plan étalé est le temps t du mouvement sur la trajectoire, défini par la relation $dt = dx/y$ (le temps est également appelé intégrale elliptique *de première espèce*).

Exercice. Supposons que l'énergie potentielle est un polynôme de degré 4 présentant deux minimums. Montrer que les périodes d'oscillations (pas forcément petites) de même énergie totale coïncident dans les deux fosses.

Indication. Les intégrales de première espèce sont identiques sur deux méridiens quelconques du tore.

Exercice. Supposons que l'énergie potentielle est un polynôme du troisième degré présentant un maximum et un minimum locaux. Montrer que la période des oscillations dans la fosse est égale à celle du mouvement de l'infini à l'infini sur une trajectoire non compacte avec la même valeur de l'énergie totale.

Remarque. Avec une énergie potentielle du troisième ou du quatrième degré on peut obtenir n'importe quelle courbe elliptique. Donc les résultats des exercices précédents entraînent que la courbe elliptique est une variété algébrique.

B. Fibrés simples au-dessus d'une courbe elliptique.

Une surface élémentaire contenant une courbe elliptique est le produit cartésien d'une courbe elliptique par la droite complexe. De même qu'au-

*) La résolution de cet exercice et des suivants fait appel aux notions élémentaires de topologie des surfaces riemanniennes figurant dans tout cours de théorie des fonctions de variable complexe.

pessus d'un cercle, outre le produit direct de ce cercle par une droite, il existe un fibré non trivial ayant une droite pour fibre (ruban de Möbius), de même au-dessus d'une courbe elliptique, outre le produit direct, il existe d'autres fibrés de fibre $\mathbb{C}$.

Fibrons le plan de deux variables complexes en droites complexes. Les fibres de ce fibré seront appelées droites *verticales*.

On désignera les coordonnées de $\mathbb{C}^2$ par (r, φ), r étant assimilé à la coordonnée verticale, φ à la coordonnée horizontale. Le fibré $\mathbb{C}^2 \to \mathbb{C}$ associe au point (r, φ) le point φ de la droite complexe horizontale.

Soit Γ une courbe elliptique recouverte par la droite horizontale. La courbe Γ s'obtient à partir de l'axe horizontal φ par identification des points différant de multiples entiers des périodes (ω_1, ω_2).

Identifions sur le plan $\mathbb{C}^2$ les droites verticales dont les projections sur une horizontale diffèrent de multiples entiers de périodes. Une telle identification transforme $\mathbb{C}^2$ en un fibré au-dessus de la courbe elliptique Γ. Or cette identification des droites verticales peut être effectuée de diverses manières (de même que lorsqu'on colle un fibré au-dessus d'un cercle, à partir d'un rectangle on peut obtenir soit un cylindre soit un ruban de Möbius selon le procédé de collage des droites verticales).

Le procédé de collage le plus simple consiste à identifier le point (r, φ) aux points $(r, \varphi + \omega_1)$ et $(r, \varphi + \omega_2)$. On obtient ainsi un produit direct. Dans le second (par ordre de difficulté) procédé de collage les droites verticales collées subissent une rotation.

E x e m p l e. Soient λ un nombre complexe différent de zéro, Γ une courbe elliptique de période $(2\pi, \omega)$. Identifions sur le plan $\mathbb{C}^2$ rapporté aux coordonnées (r, φ) les points

$$(r, \varphi), \ (r, \varphi + 2\pi), \ (\lambda r, \varphi + \omega).$$

Cette identification transforme $\mathbb{C}^2$ en une surface complexe différentiable Σ et le fibré $\mathbb{C}^2 \to \mathbb{C}$, $(r, \varphi) \mapsto \varphi$ en le fibré $\Sigma \to \Gamma$, dont la base est la courbe elliptique Γ, et de fibre $\mathbb{C}$. L'équation $r = 0$ définit un plongement de Γ dans Σ.

On peut se représenter la surface Σ comme suit (pour λ réel). Considérons un espace réel à trois dimensions muni d'un plan horizontal $\{\varphi \in \mathbb{C}\}$ fibré en droites verticales. Soit la bande $0 \leqslant \operatorname{Im} \varphi \leqslant \operatorname{Im} \omega$. Collons les plans verticaux limitant cette bande en identifiant (r, φ) du plan vertical $\operatorname{Im} \varphi = 0$ (r est la coordonnée sur la verticale) au point $(\lambda r, \varphi + \omega)$ du plan $\operatorname{Im} \varphi = \operatorname{Im} \omega$. Collons ensuite les points différant seulement de 2π sur φ. On obtient un fibré au-dessus de la courbe elliptique dont la fibre est une droite.

Pour se représenter la surface complexe Σ il reste à remplacer les droites verticales réelles par des droites verticales complexes.

Topologiquement le fibré construit est un produit direct. Cependant du point de vue holomorphe ce fibré n'est génériquement pas trivial.

C. Fibrés triviaux et non triviaux.

T h é o r è m e. *Supposons que* $\lambda \neq e^{ik\omega}$, $k \in \mathbb{Z}$. *Alors aucun voisinage de la courbe elliptique* Γ *dans la surface* Σ *décrite plus haut ne s'applique biholomorphement sur un voisinage de cette courbe* Γ *dans le produit direct.*

◀ Dans le produit direct on peut déformer la courbe Γ : pour tout ε l'équation $r = \varepsilon$ définit une courbe elliptique sur le produit direct. Soit Γ_1 une courbe elliptique sur le fibré Σ proche de la courbe Γ, section nulle du fibré (l'équation de Γ est $r = 0$). Alors la courbe Γ_1 est donnée par l'équation $r = f(\varphi)$, où $f(\varphi + 2\pi) \equiv f(\varphi)$, $f(\varphi + \omega) \equiv \lambda f(\varphi)$. En développant f en série de Fourier : $f = \Sigma f_k e^{ik\varphi}$, on trouve $f_k e^{ik\omega} = \lambda f_k$. Donc, $f_k = 0$ et Γ_1 coïncide avec Γ. Par suite, notre courbe elliptique n'est pas déformable dans le fibré avec $\lambda \neq e^{ik\omega}$. ▶

E x e r c i c e. Montrer que pour $\lambda = e^{ik\omega}$ le fibré $\Sigma \to \Gamma$ est un produit direct (holomorphement trivial).

E x e r c i c e. Montrer que les fibrés $\Sigma_1 \to \Gamma$, $\Sigma_2 \to \Gamma$ définis par les nombres complexes λ_1, λ_2 sont biholomorphement équivalents si et seulement si $\lambda_1 = \lambda_2 e^{ik\omega}$ pour un k entier.

R e m a r q u e. Les classes d'équivalence biholomorphe des fibrés décrits au-dessus d'une courbe elliptique fixe Γ constituent un groupe (la multiplication consiste à multiplier les nombres λ).

Des résultats des exercices précédents il suit que ce groupe s'identifie naturellement au groupe quotient du groupe multiplicatif des nombres complexes par le sous-groupe des nombres de la forme $e^{ik\omega}$. Le groupe quotient $\mathbb{C}^*/\{e^{ik\omega}\}$ s'applique lui-même biholomorphement sur la courbe elliptique initiale. Ce groupe s'appelle aussi *groupe de Picard* ou *variété de Jacobi* de la courbe Γ (ces notions sont définies aussi bien pour les courbes elliptiques que pour les variétés algébriques arbitraires et dans le cas général ne coïncident pas avec la variété initiale).

E x e r c i c e. Soit le fibré au-dessus d'une courbe elliptique, défini par les identifications

$$(r, \varphi) \sim (\lambda_1 r, \varphi + \omega_1) \sim (\lambda_2 r, \varphi + \omega_2).$$

Prouver que ce fibré est biholomorphement équivalent au fibré pour lequel $\omega_1 = 2\pi$, $\lambda_1 = 1$.

R e m a r q u e. On démontre que tous les fibrés vectoriels topologiquement triviaux à une dimension au-dessus d'une courbe elliptique sont biholomorphement équivalents aux fibrés $\Sigma \to \Gamma$ décrits plus haut.

D. Fibrés topologiquement non triviaux au-dessus d'une courbe elliptique. Topologiquement tous les fibrés décrits plus haut sont triviaux (homéomorphes à un produit direct). L'indice de self-intersection de la section nulle avec elle-même est un invariant qui permet de distinguer les fibrés topologiquement non équivalents.

Soient M_1, M_2 des sous-variétés compactes, orientées, différentiables d'une variété différentiable réelle orientée M (toutes ces variétés sont sans bord). Supposons que la dimension de M est égale à la somme de celles de M_1 et M_2, que M_1 et M_2 se coupent transversalement (c'est-à-dire en chaque point d'intersection la somme des espaces tangents aux deux variétés est un espace tangent à la variété ambiante M).

On appelle *indice d'intersection* de M_1 et M_2 dans M le nombre de points d'intersection signé en fonction des orientations (un point d'intersection est affecté du signe plus si le repère orientant positivement M_1 et le repère suivant orientant positivement M_2 définissent un repère orientant positivement M).

Supposons que M_1 est une sous-variété compacte différentiable orientée, de dimension égale à la moitié de celle de M. L'*indice de self-intersection* de M_1 dans M est défini comme l'indice d'intersection de M_1 et d'une variété M_2 déduite de M_1 par une petite déformation et coupant M_1 transversalement. Par exemple, l'indice de self-intersection d'un méridien du tore est nul car les méridiens voisins ne se coupent pas.

On montre que l'indice de self-intersection de M_1 dans M ne dépend pas du choix de la variété M_2 pourvu qu'elle se déduise de M_1 par une petite déformation.

E x e r c i c e. Trouver l'indice de self-intersection d'une sphère S^2 dans son fibré tangent.

R é p o n s e. $+2$. En général l'indice de self-intersection d'une variété dans son fibré tangent est égal à la caractéristique d'Euler de cette variété.

Etudions maintenant un fibré vectoriel à une dimension $\Sigma \to \Gamma$ au-dessus de la courbe elliptique obtenue à partir du fibré $\mathbb{C}^2 \to \mathbb{C}$ par collage des droites verticales (la coordonnée verticale est désignée par r) d'après la règle

$$(r, \varphi) \sim (r, \varphi + 2\pi) \sim (\lambda e^{ip\varphi} r, \varphi + \omega),$$

où p est un entier, λ un nombre complexe non nul.

Exercice. Trouver l'indice de self-intersection de la section nulle ($r = 0$) dans le fibré obtenu en admettant que Σ est orientée comme une variété complexe (l'orientation des variétés complexes est définie de telle sorte que l'indice d'intersection des plans complexes est toujours positif : l'espace muni des coordonnées complexes $(z_1, \ldots, z_n)$ est orienté par les coordonnées (Re z_1, Im z_1, . . ., Re z_n, Im z_n).

Réponse. $-p$ si Im $\omega > 0$.

Remarque. Les fibrés obtenus épuisent à une équivalence biholomorphe près tous les fibrés vectoriels à une dimension au-dessus d'une courbe elliptique.

E. Voisinage d'une courbe elliptique sur une surface complexe.

Etudions une courbe elliptique Γ sur une surface complexe Σ. Un voisinage de la courbe Γ sur la surface Σ définit un fibré vectoriel à une dimension au-dessus de Γ appelé *fibré normal.* La fibre du fibré normal en un point de Γ est l'espace quotient de l'espace tangent à Σ en ce point par le sous-espace tangent à Γ.

Le fibré normal est lui-même surface complexe. La courbe elliptique initiale Γ est plongée dans cette surface (comme section nulle du fibré).

Il se pose la question de savoir si un voisinage assez petit de Γ sur la surface initiale Σ s'applique biholomorphement sur son voisinage dans le fibré normal. Cette question est assez proche de la réductibilité d'une équation différentielle (ou d'une application différentiable) au voisinage du point fixe à la forme normale linéaire et se résout par les mêmes méthodes.

Montrons d'abord que le voisinage de la courbe elliptique Γ sur la surface n'admet pas de fibration holomorphe au-dessus de Γ.

Exemple. Soit une famille de courbes elliptiques telles que des courbes voisines de cette famille ne soient pas biholomorphement équivalentes l'une à l'autre. On obtient une telle famille, par exemple, en identifiant sur le plan des deux variables complexes (φ, ω) les points (φ, ω), $(\varphi + 2\pi, \omega)$, $(\varphi + \omega, \omega)$. Le domaine Im $\omega > 0$ se transforme par ces identifications en la réunion des courbes elliptiques $\omega =$ const. Aucune de ces courbes ne possède un voisinage s'appliquant holomorphement sur elle de telle sorte qu'elle reste invariante.

En effet, si une telle application existait, on obtiendrait une application biholomorphe proche de l'identité des courbes elliptiques avec des ω distincts proches, ce qui est impossible.

L'exemple que nous venons de traiter est dans un certain sens exceptionnel : le voisinage de la courbe elliptique plongée dans une surface complexe avec un indice de self-intersection nul est « généralement » biholomorphement équivalent à un voisinage de la courbe dans le fibré normal (au sens dans lequel une équation différentielle est généralement équivalente à une équation linéaire au voisinage d'un point singulier). Le caractère exceptionnel de l'exemple considéré est lié au fait que le fibré normal de chacune des courbes elliptiques de la famille est trivial (est un produit direct).

F. Forme normale préliminaire.

On peut obtenir une courbe elliptique à partir d'une couronne circulaire par collage holomorphe des cercles frontières. On peut de même obtenir un voisinage d'une courbe elliptique sur une surface Σ à partir d'un voisinage de la couronne sur une autre surface par un collage holomorphe des variétés frontières. Ces variétés frontières ont une dimension réelle égale à trois ; le collage est prolongeable holomorphement au voisinage de la frontière.

Il se trouve qu'un voisinage assez petit de l'image biholomorphe d'une couronne circulaire fermée sur une surface complexe s'applique toujours biholomorphement sur un voisinage d'une couronne circulaire plongée dans la droite complexe $\mathbb{C}$ du produit direct $\mathbb{C} \times \mathbb{C}$.

Comme pour les résultats relatifs à la classification holomorphe des fibrés vectoriels à une dimension au-dessus d'une courbe elliptique, la démonstration

du résultat concernant le voisinage de la couronne n'est pas simple: elle fait appel à la théorie des fonctions de plusieurs variables complexes (faisceaux, équations elliptiques à dérivées partielles ou quelque chose de ce genre).

Nous ne nous attarderons pas sur la démonstration et nous supposerons directement que la surface qui contient la courbe elliptique se déduit par collage à partir d'un voisinage de la couronne circulaire dans le produit direct.

Nous considérons donc une surface dont les points se déduisent à partir des points (r, φ) d'un plan de deux variables complexes par les collages

$$\begin{pmatrix} r \\ \varphi \end{pmatrix} \sim \begin{pmatrix} r \\ \varphi+2\pi \end{pmatrix} \sim \begin{pmatrix} rA(r, \varphi) \\ \varphi+\omega+rB(r, \varphi) \end{pmatrix},$$

où les fonctions A et B 2π-périodiques en φ sont holomorphes au voisinage de l'axe réel φ.

La couronne circulaire est obtenue à partir de la bande $0 \leqslant \operatorname{Im} \varphi \leqslant \operatorname{Im} \omega$ par collage des points $(0, \varphi) \sim (0, \varphi + 2\pi)$ sur l'axe complexe φ; r et φ sont les coordonnées sur le produit direct.

Le couple de fonctions (A, B) définissant le collage définit un voisinage. Les fonctions A et B peuvent prendre une autre forme par un choix convenable des coordonnées (r, φ). Essayons de choisir ces coordonnées de sorte à simplifier au possible A et B.

Considérons tout d'abord le changement de coordonnées linéaire $r = C(\varphi)r$, où la fonction C est holomorphe dans la bande nouvelle, $0 \leqslant \operatorname{Im} \varphi \leqslant \operatorname{Im} \omega$ de l'axe φ, de période 2π et différente de zéro sur cette bande.

T h é o r è m e. *La fonction C définissant le changement linéaire de la coordonnée verticale peut être choisie telle que dans les nouvelles coordonnées la fonction $A(0, \varphi)$ qui figure dans l'écriture du collage soit de la forme $\lambda e^{ip\varphi}$ (p est un entier égal à l'indice de self-intersection de la courbe elliptique $r = 0$ dans la surface considérée, affecté du signe moins).*

◀ La fonction $A(0, \varphi)$ définit le fibré normal de la courbe $r = 0$ dans notre surface. Ce fibré est biholomorphement équivalent au fibré obtenu par le collage

$$(r, \varphi) \sim (r, \varphi+2\pi) \sim (\lambda e^{ip\varphi} r, \varphi+\omega)$$

(voir remarque du point E). Le changement linéaire de la coordonnée r, qui ramène le collage du fibré normal à cette forme canonique, nous conduit à la forme de la fonction $A(0, \varphi)$, indiquée plus haut. ▶

D é f i n i t i o n. On appelle *forme normale préliminaire* d'un voisinage d'une courbe elliptique sur une surface, où cette courbe possède un indice de self-intersection nul, le collage

$$\begin{pmatrix} r \\ \varphi \end{pmatrix} \sim \begin{pmatrix} r\lambda(1+ra(r, \varphi)) \\ \varphi+\omega+rb(r, \varphi) \end{pmatrix} \sim \begin{pmatrix} r \\ \varphi+2\pi \end{pmatrix},$$

où a et b sont des fonctions holomorphes au voisinage de l'axe réel φ, de période 2π en φ, λ est un nombre complexe différant de zéro.

Dans la suite, nous omettrons d'indiquer l'identification des points distincts seulement de 2π sur la coordonnée φ, autrement dit, on admettra que les fonctions sont 2π-périodiques et que la coordonnée φ appartient au cylindre $\mathbb{C} \bmod 2\pi$.

G. Forme normale formelle.

D é f i n i t i o n. Un couple de nombres (λ, ω) est résonnant si $\lambda^n = e^{ik\omega}$ pour des n et k entiers non simultanément nuls.

T h é o r è m e. *Les couples résonnants forment un ensemble partout dense dans l'espace de tous les couples de nombres complexes.*

◀ Ceci est une conséquence du fait que l'ensemble des points de la forme $i\left(\frac{k}{n}\omega+\frac{m}{n}2\pi\right)$ (k et m sont entiers, n naturel) est partout dense dans la droite complexe. ▶

T h é o r è m e. *Un couple* (λ, ω) *est résonnant si et seulement si le fibré défini par le collage*

$$(r, \varphi) \sim (r, \varphi+2\pi) \sim (\lambda r, \varphi+\omega)$$

est trivial au-dessus d'un recouvrement à n feuillets cyclique de la courbe elliptique initiale.

◀ Si $\lambda^n = e^{ik\varphi}$, alors $(r, \varphi) \sim (e^{ik\omega} r, \varphi+n\omega)$ et, par suite, le fibré au-dessus de $\mathbb{C}/2\pi\mathbb{Z}+n\omega\mathbb{Z}$ est trivial (cf. point C). La réciproque se démontre de façon analogue. ▶

D é f i n i t i o n. On appelle *collage formel* l' « application

$$f\begin{pmatrix} r \\ \varphi \end{pmatrix} = \begin{pmatrix} rA(r, \varphi) \\ \varphi+\omega+rB(r, \varphi) \end{pmatrix},$$

où A et B sont des séries entières formelles sur r, de coefficients analytiques sur l'axe réel φ, 2π-périodiques en φ, $A(0, \varphi) \neq 0$.

On appelle *changement de variables formel* l'« application »

$$g\begin{pmatrix} r \\ \varphi \end{pmatrix} = \begin{pmatrix} rC(r, \varphi) \\ \varphi+rD(r, \varphi) \end{pmatrix},$$

où C et D sont des séries entières formelles sur r de coefficients 2π-périodiques en φ, analytiques dans la bande $0 \leqslant \operatorname{Im}\varphi \leqslant \operatorname{Im}\omega$ de l'axe complexe φ, $C(0, \varphi) \neq 0$.

Le changement de variables formel g opère sur le collage formel f d'après la formule $f \longmapsto g \circ f \circ g^{-1}$ (le second membre est déterminé par substitution naturelle des séries entières et lui-même est un collage formel).

T h é o r è m e. *Si le couple* (λ, ω) *n'est pas résonnant, tout collage formel*

$$\begin{pmatrix} r \\ \varphi \end{pmatrix} \longmapsto \begin{pmatrix} r\lambda(1+ra(r, \varphi)) \\ \varphi+\omega+rb(r, \varphi) \end{pmatrix}$$

se ramène par un changement formel de variables à la forme normale linéaire

$$\begin{pmatrix} r \\ \varphi \end{pmatrix} \longmapsto \begin{pmatrix} \lambda r \\ \varphi+\omega \end{pmatrix}.$$

◀ Nous allons éliminer successivement les termes d'ordre 1, 2, . . . en r dans ra et rb. Pour cela il nous faut comme toujours résoudre une équation homologique linéaire. Ecrivons l'équation de la normalisation des termes d'ordre n.

L e m m e. *Soit l'équation en u*

$$\lambda^n u(\varphi+\omega) - u(\varphi) = v(\varphi),$$

où v est une fonction 2π*-périodique analytique dans la bande* $\alpha \leqslant \operatorname{Im}\varphi \leqslant \beta$. *Si* $\tau = \operatorname{Im}\omega > 0$, $\lambda \neq 0$ *et* $\lambda^n \neq e^{ik\omega}$ *pour k entiers, alors l'équation a une solution u* 2π*-périodique, analytique dans la bande* $\alpha \leqslant \operatorname{Im}\varphi \leqslant \beta+\tau$.

◀ Soit

$$u(\varphi) = \sum u_k e^{ik\varphi}, \quad v(\varphi) = \sum v_k e^{ik\varphi}.$$

Alors

$$u_k = \frac{v_k}{\lambda^n e^{ik\omega}-1}.$$

Pour $k \to +\infty$, $|v_k|$ est majorée par une quantité de l'ordre de $e^{k(\alpha-\varepsilon)}$, et $e^{ik\omega}$ tend vers zéro. Donc $|u_k|$ est majorée par une quantité de l'ordre de $e^{k(\alpha-\varepsilon)}$.

Pour $k \to -\infty$, $|v_k|$ est majorée par une quantité de l'ordre de $e^{-|k|(\beta+\varepsilon)}$ et $|e^{ik\omega}|$ croît comme $e^{|k|\tau}$ ($\tau = \mathrm{Im}\omega > 0$). Par suite, $|u_k|$ est majoré par $e^{-|k|(\beta+\tau+\varepsilon)}$, d'où la convergence de la série de Fourier pour u au voisinage de la bande $\alpha \leqslant \mathrm{Im}\varphi \leqslant \beta + \tau$. ▶

Supposons que $ra = r^n a_n(\varphi) + \ldots$, $rb = r^n b_n(\varphi) + \ldots$, où les points de suspension désignent les termes d'ordre $>n$.

Effectuons un changement formel de variables dans lequel $C(r, \varphi) = 1 + r^n C_n(\varphi)$, $rD(r, \varphi) = r^n D_n(\varphi)$. Par une substitution directe on s'assure que les coefficients en r^n de ra et rb deviennent

$$\widetilde{a}_n(\varphi) = a_n(\varphi) + \lambda^n C_n(\varphi+\omega) - C_n(\varphi),$$

$$\widetilde{b}_n(\varphi) = b_n(\varphi) + \lambda^n D_n(\varphi+\omega) - D_n(\varphi).$$

Tirons C_n et D_n des équations $\widetilde{a}_n = 0$ et $\widetilde{b}_n = 0$. Le lemme dit que ces équations possèdent des solutions analytiques dans la bande $0 \leqslant \mathrm{Im}\varphi \leqslant \tau$. Nous avons construit un changement formel qui accroît l'ordre des termes inférieurs dans ra et rb. En itérant cette construction pour $n = 1, 2, \ldots$ on obtient un changement formel de variables qui élimine complètement ra et rb. ▶

H. Forme normale analytique.

D é f i n i t i o n. Un couple de nombres complexes (λ, ω), où $\mathrm{Im}\omega \neq 0$, $\lambda \neq 0$, est normal si existent des constantes $C > 0$, $\nu > 0$, telles que

$$|\lambda^n e^{ik\omega} - 1| > C(|n| + |k|)^{-\nu}$$

pour tous les k et n ($n \neq 0$) entiers.

On démontre immédiatement le

T h é o r è m e. *Les couples non normaux forment un ensemble partout dense, de mesure de Lebesgue nulle pour tout ω fixe.*

T h é o r è m e. *Si (λ, ω) est un couple normal alors tout collage holomorphe*

$$\begin{pmatrix} r \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} r\lambda(1 + ra(r, \varphi)) \\ \varphi + \omega + rb(r, \varphi) \end{pmatrix}$$

se ramène à la forme normale linéaire $(r, \varphi) \mapsto (r\lambda, \varphi + \omega)$ par un changement de variables holomorphe.

◀ La démonstration est analogue à celle du théorème de Siegel (§ 28). ▶

Traduisons ce théorème dans le langage des plongements de courbes elliptiques.

D é f i n i t i o n. Un fibré vectoriel holomorphe ξ est rigide si pour tout plongement de sa base dans une variété complexe telle que le fibré normal soit ξ, un voisinage assez petit de la base plongée dans la variété s'applique biholomorphement sur un voisinage de la section nulle du fibré ξ.

Dans ces termes le théorème s'énonce comme suit.

C o r o l l a i r e. *Presque tous (au sens de la mesure de Lebesgue) les fibrés vectoriels à une dimension de degré zéro au-dessus d'une courbe elliptique sont rigides.*

R e m a r q u e. Les séries formelles réduisant le collage à la forme normale peuvent diverger pour certains fibrés non résonnants dans lesquels les couples (λ, ω) ne sont pas normaux. Ces couples forment un ensemble partout dense de mesure nulle. Cette question est traitée en détail au § 36.

I. Voisinages négatifs.

Considérons le cas où l'indice de self-intersection d'une courbe elliptique sur une surface est différent de zéro. Si cet indice est négatif la courbe n'est pas déformable dans la classe des courbes holomorphes. En effet, dans le cas contraire,

l'indice d'intersection de la courbe déformée et de la courbe initiale serait non négatif (puisque ces deux courbes sont complexes).

Donc, la courbe d'indice de self-intersection négatif est isolée dans la surface. Une telle courbe est dite courbe exclusive et son voisinage, voisinage négatif.

Théorème (de Grauert). *Le fibré normal d'une courbe exclusive est toujours rigide, c'est-à-dire un assez petit voisinage d'une courbe d'indice de self-intersection négatif sur une surface complexe est défini (à une équivalence holomorphe près) par le fibré normal de cette courbe.*

Esquissons une démonstration simple pour le cas d'une courbe elliptique.

◀ On commencera par la forme normale préliminaire du collage

$$f\begin{pmatrix} r \\ \varphi \end{pmatrix} = \begin{pmatrix} r\lambda e^{ip\varphi}(1+ra(r, \varphi)) \\ \varphi+\omega+rb(r, \varphi) \end{pmatrix}.$$

Supposons que les termes d'ordre inférieur à n en r ont déjà été éliminés, c'est-à-dire $ra = r^n a_n(\varphi) + \ldots$, $rb = r^n b_n(\varphi) + \ldots$ Soit le changement de variables formel $g(r, \varphi) = (r(1 + r^n C_n(\varphi)), \varphi + r^n D_n(\varphi))$. Les coefficients en r^n dans ra et rb s'écrivent alors (c'est-à-dire pour le collage $g \circ f \circ g^{-1}$)

$$\tilde{a}_n(\varphi) = a_n(\varphi) + \lambda^n e^{ipn\varphi} C_n(\varphi + \omega) - C_n(\varphi) - ipD_n(\varphi),$$
$$\tilde{b}_n(\varphi) = b_n(\varphi) + \lambda^n e^{ipn\varphi} D_n(\varphi + \omega) - D_n(\varphi).$$

Annulons $\tilde{b}_n$ et $\tilde{a}_n$. Il nous faut d'abord tirer D_n de la seconde équation, et ensuite C_n de la première. Dans les deux cas il faut résoudre une équation homologique de la forme

$$\lambda^n e^{ipn\varphi} u(\varphi+\omega) - u(\varphi) = v(\varphi)$$

où u est une fonction inconnue 2π-périodique, v une fonction connue 2π-périodique.

J. Etude de l'équation homologique.

Considérons les développements des fonctions u et v en séries de Fourier

$$u = \sum u_k e^{ik\varphi}, \qquad v = \sum v_k e^{ik\varphi}.$$

Pour les coefficients de Fourier on obtient les équations

$$\lambda^n e^{i(k-pn)\omega} u_{k-pn} - u_k = v_k.$$

Ces équations permettent en principe de calculer successivement tous les coefficients inconnus u_k en fonction des pn premiers d'entre eux. Cependant les séries formelles de Fourier obtenues ne sont pas toujours convergentes. Or la négativité de l'indice modal de la courbe elliptique initiale sur la surface (c'est-à-dire la positivité du nombre p) implique la convergence.

En effet, soit d'abord l'équation homogène, c'est-à-dire on admettra que tous les v_k sont nuls.

Notre équation établit une relation entre les valeurs de u_k, k étant les termes d'une progression arithmétique de raison pn. Calculons toutes les valeurs de u_k en fonction de l'une d'elles. Il nous faut multiplier successivement des nombres de la forme $\lambda^n e^{i(k-pn)\omega}$. Les logarithmes de ces nombres forment une progression arithmétique de raison $ipn\omega$. Donc, les sommes des logarithmes forment une suite de la forme

$$\alpha s^2 + \beta s + \gamma$$

où s est l'indice d'un terme de la suite, $2\alpha = ipn\omega$.

Si $p > 0$, $\operatorname{Im}\omega > 0$ alors $\operatorname{Re}\alpha < 0$. Dans ce cas la suite $|e^{\alpha s^2+\beta s+\gamma}|$ tend rapidement vers zéro pour $s \to +\infty$ et $s \to -\infty$. D'où il suit que l'équation homologique homogène possède pour $p > 0$ pn solutions linéairement indépendantes décroissant rapidement pour $|k| \to \infty$.

Penchons-nous maintenant sur l'équation non homogène. Supposons tout d'abord que seul un des coefficients de Fourier de la fonction connue, soit v_m, est différent de zéro. A gauche de m on posera $u_k = 0$, et pour $k \geqslant m$ on définira les u_k à partir de l'équation. Donc, à droite de m les u_k seront confondus avec l'une des solutions de l'équation homogène et, par suite, décroîtront comme $|e^{\alpha s^2}|$.

Dans le cas général, la solution de l'équation homogène se construit comme une combinaison linéaire des solutions construites avec les coefficients v_k. La convergence est assurée par la condition $\operatorname{Re}\alpha < 0$, c'est-à-dire la négativité de l'indice de self-intersection de la courbe elliptique initiale sur la surface.

Si l'on effectue soigneusement les majorations mentionnées, on s'assure que l'équation homologique admet une solution dans le cas d'un indice de self-intersection négatif (c'est-à-dire dans le cas où p est positif). Ceci démontre la rigidité formelle du fibré normal négatif de la courbe elliptique sur la surface. Une analyse plus poussée de notre construction démontre la rigidité analytique (c'est-à-dire le théorème de Grauert) : la démonstration de la convergence est ici plus facile que dans le cas $p = 0$ traité aux points G et H, au même titre que le théorème de Poincaré est plus facile que celui de Siegel (§ 28). ▶

K. Voisinages positifs.

Supposons que l'indice de self-intersection de la courbe elliptique sur la surface est positif. Dans ce cas l'équation homologique étudiée au point précédent est généralement insoluble, puisque $|e^{\alpha s^2+\beta s+\gamma}|$ croît pour $|s| \to \infty$. Cela signifie que le voisinage de la courbe elliptique d'indice de self-intersection positif sur la surface complexe générale non seulement ne s'applique pas biholomorphement sur le voisinage de cette courbe dans le fibré normal, mais encore que cette application est impossible déjà au niveau des 2-jets (en négligeant les termes d'ordre trois par rapport à la distance à la courbe). Le voisinage de la courbe elliptique d'indice de self-intersection positif est dit *positif*.

D'après ce qui précède le voisinage positif de la courbe elliptique doit avoir des modules fonctionnels : la « forme normale » du voisinage doit contenir des fonctions arbitraires (voire probablement des fonctions de deux variables ou des germes de fonctions de deux variables en plusieurs points).

Alors que la courbe à indice de self-intersection négatif est individuelle sur la surface, la courbe elliptique à indice de self-intersection positif peut toujours être déformée.

Théorème (cas particulier du théorème de Riemann-Roch). *Si l'indice de self-intersection d'une courbe elliptique sur une surface est égal à p, le fibré normal possède p sections linéairement indépendantes.*

◀ Le problème se ramène à une équation homologique de la forme

$$u(\varphi+\omega) = \lambda e^{ip\varphi} u(\varphi),$$

possédant, comme nous l'avons vu au point J, p solutions linéairement indépendantes. ▶

En passant aux termes d'ordre supérieur par rapport à la distance à la courbe elliptique on s'assure qu'il existe une déformation p-paramétrique de la courbe dans son voisinage.

D'où il résulte entre autres que le voisinage d'une courbe elliptique sur une surface, où cette courbe possède un indice de self-intersection positif, en principe ne présente pas une structure de fibré au-dessus de cette courbe. En effet, la structure complexe de la courbe elliptique varie généralement par une déformation. Donc, parmi les courbes déformées voisines il existera dans le cas général des courbes ne s'appliquant pas biholomorphement sur la courbe initiale.

Quand on étudie le passage des équations différentielles par la résonance, on ne rencontre que des voisinages de courbes elliptiques sur des surfaces sur lesquelles l'indice de self-intersection est nul.

L. Courbes elliptiques dans l'espace.

Beaucoup de ce qui a été dit sur le voisinage d'une courbe elliptique sur une surface se généralise au cas d'une courbe elliptique dans un espace à plusieurs dimensions. Pour cela il faut considérer que la variable r est multidimensionnelle dans les formules citées.

Les fibrés vectoriels de dimension arbitraire au-dessus d'une courbe elliptique sont décrits par les collages

$$(r, \varphi) \sim (r, \varphi + 2\pi) \sim (\Lambda(\varphi) r, \varphi + \omega),$$

où $\Lambda(\varphi)$ est un opérateur linéaire dans la forme normale de Jordan dont les valeurs propres sont de la forme $\lambda e^{ip\varphi}$.

Un fibré est *négatif* (resp. *non positif*, *nul*) si tous les nombres p sont positifs (resp. *non négatifs*, *nuls*).

Supposons que le fibré normal d'une courbe elliptique est négatif. Alors tout voisinage de cette courbe sur une variété s'applique biholomorphement sur son voisinage dans le fibré normal (théorème de Grauert), c'est-à-dire le fibré normal est rigide. Dans la classe des fibrés normaux nuls la rigidité n'est violée qu'avec la probabilité zéro. La condition de résonance s'écrit $\lambda_s = \lambda^n e^{ik\omega}$, où k est un entier, $\lambda^n = \lambda_1^{n_1} \ldots \lambda_m^{n_m}$, $m = \dim \{r\}$, $n_j \geqslant 0$, $\sum n_j \geqslant 2$.

Un fibré non résonnant est formellement rigide. Pour qu'il soit véritablement rigide il suffit que soit réalisée l'inégalité habituelle de (C, ν)-normalité:

$$|\lambda^n e^{ik\omega} - \lambda_s| \geqslant C(|n| + |k|)^{-\nu}, \quad |n| = n_1 + \ldots + n_m,$$

pour tous les $n \geqslant 0$, $\sum n_i \geqslant 2$ et les k entiers. L'ensemble des vecteurs λ qui ne sont (C, ν)-normaux pour aucun $(C, \nu) > 0$ est de mesure nulle.

Les fibrés non positifs sont (probablement) presque tous rigides.

La construction de voisinages de courbes de genre plus grand que 1 est un problème peu étudié à l'exception du cas où le fibré normal est négatif et, par suite, rigide d'après le théorème de Grauert.

§ 28. Démonstration du théorème de Siegel

Dans ce paragraphe on démontre le théorème de l'équivalence locale holomorphe d'une application et de sa partie linéaire en son point fixe.

A. Formulation du théorème.

D é f i n i t i o n. Une collection $(\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^n$ est de *type multiplicatif* (C, ν) si

$$|\lambda_s - \lambda^k| \geqslant C|k|^{-\nu} \; (|k| = k_1 + \ldots + k_n, \; \lambda^k = \lambda_1^{k_1} \ldots \lambda_n^{k_n})$$

pour tout s et tout vecteur entier k de composantes non négatives, tel que $|k| > 1$ $(C > 0, \nu > 0)$.

T h é o r è m e. *Supposons qu'une collection de valeurs propres de la partie linéaire d'une application holomorphe de point fixe $O \in \mathbb{C}^n$ est de type multiplicatif (C, ν) au voisinage de O. Alors cette application est biholomorphement équivalente à sa partie linéaire dans un voisinage du point O.*

Supposons que A est une application biholomorphe définie au voisinage d'un point $O \in \mathbb{C}^n$ laissant le point O invariant et l'opérateur linéaire Λ est la partie linéaire de A en O. On affirme qu'il existe un difféomorphisme H biholomorphe dans un voisinage de O laissant le point O invariant et tel que $H \circ A \circ H^{-1} = \Lambda$ dans un voisinage du point O.

Nous allons prouver ce théorème pour le cas où toutes les valeurs propres λ_s de l'opérateur Λ sont distinctes. Dans ce cas on peut choisir un système de coordonnées dans lequel la matrice de l'opérateur Λ est diagonale. Fixons un tel système de coordonnées.

B. Construction du changement de coordonnées H.

Mettons l'application donnée A et le changement H sous la forme

$$A(z) = \Lambda z + a(z),\quad H(z) = z + h(z),$$

où les séries de Taylor de a et de h ne contiennent pas de termes d'ordre 0 et 1 au point O. Calculons les termes d'ordre 0 et 1 en h et a de l'application $H \circ A \circ H^{-1}$. On obtient

$$(H \circ A \circ H^{-1})(z) = \Lambda z + [a(z) - \Lambda h(z) + h(\Lambda z) + R([a], [h])(z),$$

où le résidu R est du deuxième ordre de petitesse en a et h dans un sens qui sera précisé plus bas. Nous avons compris les arguments de R dans des crochets pour souligner que cet opérateur opère sur des fonctions et non sur leurs valeurs au point z.

Etudions l'équation homologique en h

$$\Lambda h(z) - h(\Lambda z) = a(z).$$

Les séries de Taylor de la fonction vectorielle connue a et de la fonction inconnue h ne possèdent pas de termes linéaires et de termes libres. Cette équation admet une solution unique dans la classe de ces séries, puisque la collection des valeurs propres n'est pas résonnante.

Nous montrerons au point C que les séries obtenues sont convergentes si la collection de valeurs propres est de type multiplicatif (C, ν) pour des C et ν positifs. Désignons par U l'opérateur qui associe le second membre a de l'équation homologique à sa solution $h = U([a])$.

Calculons par récurrence les fonctions a_s, h_s

$$h_s = U([a_s]),\quad a_{s+1} = R([a_s], [h_s]),$$

en commençant par $a_0 = a$.

Construisons les applications $H_0, H_1, \ldots$, définies par

$$H_s(z) = z + h_s(z).$$

On verra que le changement de coordonnées cherché est donné par la formule

$$H = \lim_{s\to\infty} H_s \circ \ldots \circ H_1 \circ H_0.$$

C. Etude de l'équation homologique.

On suppose que les séries de Taylor du second membre et les solutions de l'équation homologique

$$\Lambda h(z) - h(\Lambda z) = a(z)$$

ne possèdent pas de termes libres et linéaires. On pose $|z| = \max |z_j|$.

Lemme 1. *Supposons qu'une collection de valeurs propres de l'opérateur linéaire diagonal Λ est de type multiplicatif (C, ν). Supposons que le second membre a de l'équation homologique est continu dans le polycylindre $|z_j| \leqslant r$ et holomorphe à l'intérieur de ce polycylindre.*

Sous ces conditions l'équation homologique admet une solution h holomorphe à l'intérieur du polycylindre et telle que pour tout $\delta \in [0, 1/2[$ l'on ait

$$\max_{|z|\leqslant re^{-\delta}} |h(z)| \leqslant \max_{|z|\leqslant r} |a(z)|/\delta^{\alpha},$$

où la constante $\alpha = \alpha(\Lambda)$ ne dépend ni de δ ni de a ni de r.

◀ Développons a et h en séries de Taylor et désignons respectivement par a_k^s et h_k^s les coefficients en $z^k e_s$. Alors $h_k^s = a_k^s/(\lambda_s - \lambda^k)$. Evaluons les numérateurs à l'aide des inégalités de Cauchy pour les coefficients de Taylor, et les dénominateurs en tenant compte de ce que $\{\lambda_s\}$ sont du type (C, ν).

Posons $M = \max\limits_{|z|\leqslant r} |a(z)|$. L'inégalité de Cauchy nous dit que $|a_k^s| \leqslant$ $\leqslant M/r^{|k|}$. Donc,

$$|h_k^s| \leqslant MC^{-1} |k|^\nu / r^{|k|}.$$

Evaluons la somme de Taylor $\sum h_k^s z^k$. Considérons les termes de degré $|k| = p$. Leur nombre est $\leqslant c_1(n)\, p^{n-1}$, donc

$$\Big| \sum_{|k|=p} h_k^s z^k \Big| \leqslant M c_2 p^m |(z/r)^k|,$$

où $c_2 = c_1/C$ et $m = \nu + n - 1$.

La fonction $x^m e^{-x}$ a pour maximum $(m/e)^m$, donc

$$p^m e^{-\delta p/2} \leqslant c_3 \delta^{-m},$$

où

$$c_3 = (2m/e)^m.$$

Donc, pour $|z| \leqslant re^{-\delta}$ on a

$$|h| \leqslant M c_3 \delta^{-m} \sum_{p=2}^{\infty} e^{-\frac{p\delta}{2}} = M c_3 \delta^{-m} / (e^\delta - e^{\delta/2}),$$

$$|h| \leqslant M c_4 \delta^{-(m+1)}$$

où $c_4 = 4c_2c_3$ ne dépend pas de a, r et δ. ▶

On aura encore besoin de la majoration de la fonction h définie par la condtion $(h \circ \Lambda)(z) = h(\Lambda z)$.

L e m m e 2. *Dans les conditions du lemme* 1

$$\max_{|z|\leqslant re^{-\delta}} |h(\Lambda z)| \leqslant \max_{|z|\leqslant r} \|a(z)| / \delta^{\alpha_0},$$

où la constante positive $\alpha_0 = \alpha_0(\Lambda)$ *ne dépend pas de* δ, a, r.

◀ Commençons par la remarque suivante.

Si $\{\lambda\}$ *est de type multiplicatif* (C, ν), *alors il existe une constante* C_0 *ne dépendant pas de* k, *telle que*

$$|\lambda_s - \lambda^k| \geqslant C_0 |k|^{-\nu} |\lambda^k|$$

pour tous les $s = 1, \ldots, n$ *et pour tous les vecteurs* k *entiers à composantes non négatives, dont la somme* $|k|$ *est* $\geqslant 2$.

◀ Soit $\mu = \max |\lambda_s|$. Pour $|\lambda^k| \leqslant 2\mu$ on peut prendre $C_0 = C/2\mu$. Pour $|\lambda^k| \geqslant 2\mu$ on peut prendre $C_0 = 1/2$. ▶

La proposition prouvée nous donne la majoration

$$|h(\Lambda z)| \leqslant \sum MC_0^{-1} |k|^\nu |\lambda^k|^{-1} |\lambda^k z^k / r^k|.$$

Pour les majorations suivantes voir la démonstration du lemme 1. ▶

D. Ordre des opérateurs.

Pour établir les majorations suivantes il est commode d'adopter les notations:

f est une fonction continue dans le polycylindre $|z| \leqslant r$, holomorphe en les points intérieurs et nulle au centre;

$$\| f \|_r = \sup_{0<|z|\leqslant r} |f(z)|/|z|$$

est la norme pour de telles fonctions.

Exemple. La fonction $f(z) = \varepsilon z$ a pour norme $|\varepsilon|$ indépendamment du rayon du polycylindre.

Remarque. La commodité de la norme adoptée tient à son invariance par les changements d'échelle: pour tout coefficient de dilatation $\varkappa$

$$\| \varkappa f \circ \varkappa^{-1} \|_{\varkappa r} = \| f \|_r.$$

Les valeurs de la fonction f peuvent être aussi bien des nombres que des éléments d'un espace normé, par exemple, des vecteurs des matrices, etc.

Soit Φ un opérateur sur la classe des fonctions décrites ci-dessus *). Supposons que d, α et β sont des nombres positifs et que $0 < r < 1$.

Définition. *L'opérateur* Φ *est d'ordre* $(d;\ \alpha \mid \beta)$ si $\forall \delta \in]0,\ 1/2[$ et $\forall\ r \in]0,\ 1[$

$$\| \Phi([f]) \|_{re^{-\delta}} \leqslant \| f \|_r^d\, \delta^{-\alpha},$$

pourvu que $\| f \|_r \leqslant \delta^{\beta}$.

Cette relation sera notée $\Phi([f]) \prec f^d(\alpha \mid \beta)$, ou de façon plus condensée $\Phi([f]) \prec f^d$.

L opérateur *est d'ordre* d si existent des constantes α et β telles qu'il soit d'ordre $(d;\ \alpha \mid \beta)$ (il est essentiel que α et β ne dépendent pas de f, de $r \in]0,\ 1[$ et de $\delta \in]0,\ 1/2[$).

Exemple 1. Soit l'opérateur transformant le second membre a de l'équation homologique dans sa solution h. *Cet opérateur est d'ordre* 1 *si* $\{\lambda_s\}$ *est de type* $(C,\ \nu)$. En effet, on obtient l'inégalité requise à l'aide du lemme 1.

De façon analogue, on déduit du lemme 2 que l'opérateur qui transforme le second membre a de l'équation homologique dans la fonction $h \circ \Lambda$ est d'ordre $1: h \prec a,\ \ h \circ \Lambda \prec a$.

Exemple 2. Soit le difféomorphisme local $H: H(z) = z + h(z)$. Mettons le difféomorphisme inverse sous la forme $H^{-1}(z) = z - g(z)$.

Considérons l'opérateur G qui envoie h dans g.

L'opérateur G *est d'ordre* 1, *c'est-à-dire* $g \prec h$.

◀ On remarquera tout d abord que la majoration de Cauchy entraîne l'inégalité suivante: pour $|z| \leqslant re^{-\delta/2}$

$$|\partial h_i/\partial z_j| \leqslant \| h \|_r/(1 - e^{-\delta/2}). \tag{1}$$

Si $\| h \|_r \leqslant \delta^{\beta}$ et β est assez grand, le second membre de la dernière inégalité est aussi petit que l'on veut. Maintenant on obtient g comme la limite des itérations:

$$g_{s+1}(z) = h(z - g_s(z)),\ \ g_0 = 0.$$

La convergence pour $|z| \leqslant re^{-\delta}$ et la majoration $g \prec h$ résultent désormais des théorèmes des applications contractantes. ▶

Exemple 3. Dans les notations de l'exemple 2

$$h - g \prec h^2.$$

*) On désignera par une seule lettre les opérateurs agissant sur les fonctions des classes correspondant à des r différents, sous réserve qu'elles se confondent en tant qu'opérateurs sur l'espace des germes des fonctions au même titre qu'on ne change pas la notation du sinus lorsqu'on change de domaine de définition.

◀ En effet, par définition de la fonction g,

$$h(z) - g(z) = h(z) - h(z - g(z)).$$

En se servant de l'inégalité (1) et de la majoration précédente

$$\| g \|_{re^{-\delta}} \leqslant \| h \|_r \, \delta^{-\alpha},$$

on obtient pour $|z| \leqslant re^{-\delta}$

$$|h(z) - g(z)| \leqslant C \, \| h \|_r \, (1 - e^{-\delta/2})^{-1} \, r \, \| h \|_r \, \delta^{-\alpha}. \blacktriangleright$$

On remarquera que dans nos notations $2f \prec f$, $f^2 \prec f$; si $f_1 \prec f_2$ et $f_2 \prec f_3$, alors $f_1 \prec f_3$.

Généralisons ces notations aux opérateurs de plusieurs fonctions. Supposons que l'opérateur Ξ envoie un couple de fonctions η, ζ dans une fonction ξ. Soit φ un polynôme. On écrira

$$\xi \prec \varphi(\eta, \zeta),$$

si existent des constantes positives $(\alpha, \beta_1, \beta_2)$, telles que $\forall \delta \in]0, 1/2[$, $\forall r \in]0, 1[$ on ait

$$\| \Xi([\eta], [\zeta]) \|_{re^{-\delta}} \leqslant \varphi(\| \eta \|_r, \| \zeta \|_r) \, \delta^{-\alpha},$$

pourvu que $\| \eta \|_r \leqslant \delta^{\beta_1}$, $\| \zeta \|_r \leqslant \delta^{\beta_2}$. Les constantes α et β ne doivent pas dépendre ici de η, ζ, $r \in]0, 1[$ et $\delta \in]0, 1/2[$. Si $\eta \prec \psi(\sigma, \tau)$, alors $\xi \prec \prec \overline{\psi}(\psi(\sigma, \tau), \zeta)$.

Exemple 4. *Si l'opérateur Ξ est défini par*

$$\xi(z) = \eta(z - \zeta(z)) - \eta(z),$$

alors $\xi \prec \eta\zeta$.

◀ Cette proposition se démontre à l'aide de l'inégalité (1) que nous avons utilisée dans les exemples 2 et 3. ▶

E. Estimation du résidu.

Explicitons le résidu R défini au point B. On utilisera les notations

$$H(z) = z + h(z), \quad H^{-1}(z) = z - g(z).$$

Par définition

$$R(z) = (H \circ A \circ H^{-1})(z) - \Lambda z - [a(z) - \Lambda h(z) + h(\Lambda z)].$$

Mettons R sous la forme $R = R_1 + R_2 + R_3$, où

$$R_1(z) = \Lambda(h(z) - g(z)), \quad R_2(z) = a(z - g(z)) - a(z),$$
$$R_3(z) = h(\Lambda z - \Lambda g(z) + a(z - g(z))) - h(\Lambda z).$$

Pour faciliter l'établissement des majorations on mettra R sous la forme d'un opérateur de trois arguments a, h et $u = h \circ \Lambda$.

Soient les opérateurs

$$G, \ G([h]) = g, \quad \Xi, \ \Xi([\alpha], [g])(z) = a(z - g(z)) - a(z).$$

Dans ces notations

$$R_1([h]) = \Lambda(h - G([h])),$$
$$R_2([a], [h]) = \Xi([a], G([h])),$$
$$R_3([u], [a], [h]) = \Xi([u], [v]),$$

où $v(z) = g(z) - \Lambda^{-1} a(z - g(z))$.

L'opérateur $R_1 + R_2 + R_3$ se transforme en le résidu $R([a], [h])$ par la substitution $u = h \circ \Lambda$. Supposons que $h \prec \mathrm{id}$ (id = identité; ce qui signifie que la dérivée de h est petite).

Majoration 1. *On a*

$R_1([h]) \prec h^2$, $R_2([a], [h]) \prec ah$, $R_3([u], [a], [h]) \prec u(h + a)$.

◀ La majoration $R_1 \prec h^2$ a été établie dans l'exemple 3 du point D, l'inégalité $\Xi([a], [g]) \prec ag$, dans l'exemple 4. En vertu de l'exemple 2, on a $G([R]) \prec \prec h$, donc $R_2([a], [h]) \prec ah$.

Les majorations $g \prec h$ appliquées à v donnent $v \prec h + a$. Donc, en vertu de la majoration de l'opérateur Ξ de l'exemple 4, il vient $R_3 \prec u(h + a)$. ▶

Majoration 2. *Soit U opérateur solution de l'équation homologique. Alors l'opérateur Φ défini par $\Phi([a]) = R([a], U([a]))$, est d'ordre 2.*

◀ En vertu des lemmes 1, 2 et 3 du point C, on a $h \prec a$, $h \circ \Lambda \prec a$, où $h = U([a])$.

Une comparaison avec la majoration 1 donne

$$R_1(U([a])) \prec a^2, \quad R_2([a], U([a])) \prec a^2,$$

$$R_3(U([a]) \circ \Lambda, [a], U([a])) \prec a^2. \text{ ▶}$$

F. Convergence des approximations.

On achève de démontrer le théorème de Siegel comme la majoration du § 12.

◀ Soient les suites de nombres

$$\delta_0, \ \delta_1 = \delta_0^{3/2}, \ \delta_2 = \delta_1^{3/2}, \ \ldots;$$

$$M_0 = \delta_0^N \ M_1 = M_0^{3/2}, \ \ldots, \ M_s = M_{s-1}^{3/2} = \delta_s^N, \ \ldots;$$

$$r_0, \ r_1 = e^{-\delta_0} r_0, \ r_2 = e^{-\delta_1} r_1, \ \ldots$$

Ces suites dépendent du choix de δ_0, N et r_0.

Le nombre δ_0 est choisi aussi petit qu'il faut pour que les r_s soient plus grands que $r_0/2$. Décrivons le choix de N.

En vertu de la majoration 2 il existe des constantes α et β telles que

$$\| R([a], U([a])) \|_{re^{-\delta}} \leqslant \| a \|_r^2 \, \delta^{-\alpha},$$

pourvu que $\| a \|_r \leqslant \delta^\beta$.

Définissons $a_{s+1} = R([a_s], U([a_s]))$. Supposons que $\| a_s \|_{r_s} \leqslant M_s = \delta_s^N$. Pour $N > \beta$, l'inégalité précédente avec $\delta = \delta_s$ donne

$$\| a_{s+1} \|_{r_{s+1}} \leqslant M_s^2 \delta^{-\alpha} = \delta_s^{2N-\alpha}.$$

Si $N > 2\alpha$ le second membre est $\leqslant M_{s+1} = \delta_s^{3N/2}$. On fixera donc $N > (\beta, 2\alpha)$. Alors si $\| a_0 \|_{r_0} \leqslant M_0 = \delta_0^N$, on aura $\| a_s \|_{r_s} \leqslant M_s = \delta_s^N$ pour tous les s.

Choisissons enfin r_0. Par hypothèse, la fonction initiale $a_0 = a$ possède en l'origine des coordonnées un zéro d'ordre deux au moins. Donc, au voisinage du point $z = 0$, on a

$$| a(z) | \leqslant K | z |^2.$$

D'où il suit

$$\| a_0 \|_{r_0} \leqslant K r_0.$$

Par suite, la condition $\| a_0 \|_{r_0} \leqslant \delta_0^N$ est vérifiée pour r_0 assez petit. Fixons un tel r_0. Tous les nombres δ_s, M_s, r_s sont maintenant définis. Les inégalités $\| a_s \|_{r_s} \leqslant M_s$ sont réalisées pour tous les s. Elles entraînent les majorations pour h_s. Donc les produits $H_s \circ \ldots \circ H_1$ sont définis pour $| z | \leqslant r_0/2$ et pour $s \to \infty$ convergent vers H. Il est aisé de vérifier que pour le difféomorphisme limite on a $H \circ A \circ H^{-1} = \Lambda$. ▶

14—0743

CHAPITRE 6

THÉORIE LOCALE DES BIFURCATIONS

La bifurcation est utilisée dans un sens large pour désigner toute modification qualitative, topologique par suite d'une variation des paramètres dont dépend l'objet étudié. Ces objets peuvent être des plus divers: courbes réelles ou complexes ou surfaces, fonctions ou applications, variétés ou fibrés, champs de vecteurs ou équations différentielles ou intégrales.

Si l'objet dépend de paramètres, on dit qu'est donnée une famille. Si les paramètres varient légèrement au voisinage de valeurs fixées, on parle de déformation de l'objet correspondant à ces valeurs des paramètres.

Il se trouve que l'étude de toutes les déformations se ramène souvent à celle d'une seule d'où l'on déduit toutes les autres. Une telle déformation doit être dans un certain sens la plus riche, doit donner toutes les bifurcations possibles de l'objet considéré; on l'appelle déformation verselle.

Dans ce chapitre on étudie essentiellement les bifurcations et les déformations des trajectoires des systèmes dynamiques au voisinage des positions d'équilibre et des trajectoires fermées.

§ 29. Familles et déformations

On développe dans ce paragraphe les généralités « euristiques » qui servent de fondement à la théorie des bifurcations. Ces généralités sont dues essentiellement à H. Poincaré.

A. Cas génériques et cas singuliers de codimension petite.

En général, lorsqu'on étudie des objets analytiques de diverse nature (équations différentielles, problèmes aux limites ou problèmes d'optimisation) il est possible de distinguer un cas générique. Par exemple, parmi les points singuliers d'un champ de vecteurs sur le plan les nœuds, les foyers et les cols sont des points génériques, tandis que, disons, les centres se détruisent par une déformation aussi petite que l'on veut du champ.

L'étude des cas génériques est primordiale dans l'analyse des phénomènes et processus décrivant un modèle mathématique donné. En effet, une déformation aussi petite que l'on veut du modèle transforme le cas non générique en cas générique, quant aux paramètres du modèle ils sont définis approximativement.

Cependant il est parfois indispensable d'étudier des cas non génériques. Plus exactement, supposons que nous étudions non pas un objet particulier (un champ de vecteurs, par exemple) mais toute une famille dépendant d'un certain nombre de paramètres.

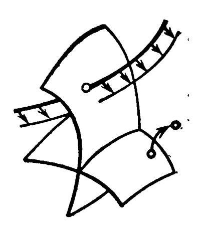


Fig. 103

Pour mieux nous représenter la situation considérons un espace fonctionnel dont les points sont les objets étudiés (par exemple, l'espace de tous les champs de vecteurs). Le cas non générique correspond à certaines hypersurfaces de codimension 1 dans cet espace. Une translation aussi petite que l'on veut transporte un point d'une hypersurface dans le domaine des cas génériques. Les hypersurfaces singulières constituent les frontières des domaines des cas génériques (fig. 103).

Une famille à k paramètres est représentée par une variété de dimension k dans notre espace fonctionnel. Une famille à un paramètre, par exemple, est une courbe de l'espace fonctionnel (représentée par une ligne grasse sur la figure).

Cette courbe est susceptible de couper une hypersurface singulière. Si cette intersection a lieu « sous un angle non nul » (transversalement), alors elle est préservée par une petite déformation de la famille: toute courbe voisine coupe cette hypersurface en un point voisin (la ligne fine sur la figure).

Donc, même si chaque membre de la famille peut être ramené au cas générique par une déformation aussi petite que l'on veut, il est impossible de rendre tous les membres de la famille simultanément génériques: par une déformation de la famille on peut éviter le cas non générique pour chaque valeur fixe du paramètre, mais ce cas se présente tout de même pour une valeur voisine du paramètre.

En général, les hypersurfaces des cas singuliers de notre espace fonctionnel présentent leurs propres singularités (par exemple, au point d'intersection de deux hypersurfaces, point qui correspond à l'apparition simultanée de deux singularités (cf. fig. 103). Lorsqu'on étudie des familles génériques à un paramètre, on peut ne pas tenir compte des singularités des hypersurfaces des cas singuliers. En effet, l'ensemble de toutes ces singularités est de codimension 2. Donc, dans l'espace fonctionnel, on peut extirper de ces singularités une courbe par une déformation aussi petite que l'on veut, de sorte que cette courbe ne coupera les hypersurfaces des cas singuliers qu'en des points génériques. Ainsi, dans une famille à un paramètre générique les seuls points singuliers inévitables sont des singularités

simples correspondant à des points non singuliers d'une hypersurface de cas singuliers. Ces singularités sont dites *singularités de codimension un*. L'étude des singularités de codimension un permet de passer continûment d'un point générique quelconque de l'espace fonctionnel à un autre point générique, puisque les ensembles de codimension >1 ne partitionnent pas l'espace fonctionnel.

En général, lors de ce passage on doit couper la surface de singularités de codimension un. L'étude des singularités de codimension un permet de décrire les bifurcations qui apparaissent sur ces surfaces.

Lorsqu'on étudie des familles à k paramètres génériques, seules les singularités de codimension $\leqslant k$ sont inévitables. Toutes les autres forment dans l'espace fonctionnel un ensemble de codimension $>k$ et on peut les éliminer par une déformation aussi petite que l'on veut de la famille à k paramètres.

Plus la singularité est de codimension élevée, plus il est difficile de l'étudier, et moins (en général) cette étude est profitable. L'étude des singularités de codimension k élevé ne présente de l'intérêt que dans le cas où elle porte non pas sur un seul objet (par exemple un champ de vecteurs avec un point singulier dégénéré), mais sur une famille à k paramètres tellement grande que la singularité du type considéré n'est pas éliminée par une petite déformation de cette famille.

Ce raisonnement simple de Poincaré montre l'inanité du nombre si grand de recherches effectuées en théorie des équations différentielles et dans d'autres domaines de l'analyse qu'on a toujours quelque appréhension à le rappeler. En fait, toute étude d'un cas dégénéré doit être conduite de pair avec les calculs de la codimension correspondante et l'indication des bifurcations dans la famille pour laquelle cette singularité est inévitable.

De ce point de vue qui est basé sur l'étude de familles à k paramètres on peut ne pas se soucier des singularités de codimension infinie, car on peut les éliminer par une petite déformation d'une famille à k paramètres, k fini. Il est entendu que les cas dégénérés peuvent être utiles en tant que premières approximations facilement étudiables de la théorie des perturbations.

B. Digression sur les cas de codimension infinie.

Parfois on a à étudier des singularités de codimension infinie. Par exemple, les systèmes hamiltoniens ou les systèmes possédant un groupe de symétrie quelconque forment une sous-variété de codimension infinie dans l'espace de tous les systèmes dynamiques. Dans ces cas, on arrive souvent à restreindre à priori l'espace fonctionnel, de telle sorte que les codimensions des singularités étudiées soient finies (par exemple, en se limitant uniquement aux systèmes hamiltoniens et à leurs déformations hamiltoniennes).

Du reste, cette restriction de l'espace fonctionnel n'est pas toujours facile à effectuer. Considérons, par exemple, les problèmes aux limites pour des équations aux dérivées partielles. On a en vue l'intersection de deux sous-variétés dans l'espace fonctionnel: l'espace des solutions et l'espace des fonctions vérifiant des conditions aux limites. Ces variétés possèdent toutes deux des dimensions infinies et des codimensions infinies. Pour analyser cette situation il faut

avoir distinguer les diverses dimensions et codimensions infinies: la condition de nullité d'une fonction d'une variable construite d'après l'objet donné définit dans l'espace fonctionnel une « variété dont la codimension (infinie) est plus petite » que celle de la variété définie par la condition de nullité d'une fonction de deux variables.

Un problème simple faisant intervenir un tel calcul des codimensions infinies correspondant aux noyaux et conoyaux composés de fonctions sur des variétés de dimension différente est le problème de la dérivée oblique qui s'énonce ainsi : sur la sphère limitant une boule à n dimensions, on donne un champ de vecteurs tangent à l'espace ambiant à n dimensions et on demande une fonction harmonique sur cette boule, dont la dérivée suivant la direction de ce champ est égale sur la frontière à une fonction donnée.

Traitons le cas $n = 3$, par exemple. Dans ce cas, le champ générique est tangent à la sphère le long d'une courbe différentiable. Cette courbe présente encore des points singuliers en lesquels le champ lui est tangent. La structure du champ au voisinage de chacun de ces points singuliers est standard : on démontre que pour tout n pour le champ générique *) le champ est donné au voisinage de chaque point de la frontière dans un système de coordonnées convenables par une formule de la forme

$$x_2\partial_1 + x_3\partial_2 + \ldots + x_k\partial_{k-1} + \partial_k, \quad k \leqslant n,$$

où $\partial_k = \partial/\partial x_k$ et $x_1 = 0$ sur la frontière (cf. S. Vichik, Sur le problème de la dérivée oblique, Vestnik MGU, ser. matem., 1 (1972), 21-28).

Le problème de la dérivée oblique doit probablement être posé d'après le schéma suivant. Les variétés de contact du champ avec le bord, les variétés de contact du champ avec les variétés de contact, etc. partagent la frontière en parties de dimensions distinctes. Sur certaines de ces parties il faut donner des conditions aux limites, sur les autres, au contraire, la fonction donnée sur la frontière doit elle-même vérifier certaines conditions pour que le problème admette une solution classique.

Malgré l'abondance des études du problème de la dérivée oblique (ce problème a été le mieux traité dans les travaux de V. Mazia ; et avant lui par M. Malioutov, Y. Egorov et V. Kondratiev) le programme décrit plus haut n'a été réalisé que dans le cas à deux dimensions (lorsque la frontière est un cercle).

C. Espaces des jets.

L'étude des bifurcations dans les familles génériques à k paramètres revient au fond à étudier un espace fonctionnel décomposé en parties correspondant aux diverses singularités, de codimension $\leqslant k$ (on néglige celles de codimension $>k$).

Pour se débarrasser de la dimension infinie de l'espace fonctionnel, on a mis au point un appareil spécial d'approximations de dimensions finies : les variétés de k-jets (ce terme est dû à Eresman).

Plus bas on introduira les termes et les notations qui seront utilisés. Les propositions de ce point et du suivant sont évidentes.

Soit $f\colon M^m \to N^n$ une application différentiable de variétés différentiables (on peut admettre que M et N sont des domaines d'espaces euclidiens de dimension respective m et n).

---

*) i.e. pour tout champ d'un ensemble ouvert partout dense dans l'espace fonctionnel des champs différentiables.

D é f i n i t i o n. Deux telles applications f_1, f_2 sont dites tangentes d'ordre k ou k-tangentes en un point x de M si (fig. 104)

$$\rho_N(f_1(y), f_2(y)) = o(\rho_M^k(x, y)) \text{ pour } y \to x;$$

ρ désigne une métrique riemannienne quelconque; il est aisé de voir que la propriété de k-tangence ne dépend pas du choix des métriques ρ_M et ρ_N.

Deux applications sont 0-tangentes en un point x si elles prennent la même valeur en ce point. La tangence d'ordre k est une relation d'équivalence ($f_1 \sim f_2 \Rightarrow f_2 \sim f_1$, $f_1 \sim f_2 \sim f_3 \Rightarrow$ $\Rightarrow f_1 \sim f_3$, $f_1 \sim f_1$).

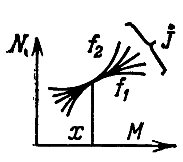


Fig. 104

D é f i n i t i o n. On appelle k-jet d'une application différentiable en un point x la classe des applications k-tangentes en x.

N o t a t i o n.

$$j_x^k(f) = \{f_1 : f_1 \text{ est } k\text{-tangent à } f \text{ en } x\}.$$

Le point x est appelé *source*, le point $f(x)$ *but* de ce jet.

Choisissons des coordonnées sur M et N au voisinage des points x et $f(x)$ respectivement. Alors un k-jet d'une application quelconque proche de f est donné en un point quelconque proche de x par une collection de coefficients d'un polynôme de Taylor de degré k. Donc, dans un système de coordonnées fixes, on peut identifier un k-jet à une collection de coefficients d'un polynôme de Taylor de degré k.

E x e m p l e. Un 0-jet d'une application f de l'axe x dans l'axe y en un point x est donné par un couple de nombres (x, y), où $y = f(x)$. Un 1-jet est donné par le triplet de nombres (x, y, p), où $p = df/dx$.

Outre la tangence d'ordre k il existe encore une relation d'équivalence qui conduit aux *germes* des applications.

D é f i n i t i o n. Deux applications définies dans deux voisinages d'un même point *possèdent un germe commun* en ce point si elles sont confondues dans un troisième voisinage de ce point (le troisième voisinage peut être plus petit que l'intersection des deux premiers).

On appelle *germe d'une application* en un point la classe d'équivalence introduite.

Tout comme pour une application on peut définir le 0-jet, le 1-jet, etc., pour un germe.

Considérons l'ensemble de tous les k-jets des germes *) des applications différentiables de M dans N en tous les points de M.

---

*) Dans le cas différentiable réel il est indifférent de considérer les jets de germes ou les jets des applications de M tout entière, car chaque germe est germe d'une application globale. Dans le cas complexe il peut ne pas exister d'application différentiable globale de jet donné.

D é f i n i t i o n. On appelle *espace des k-jets des applications de M dans N* l'ensemble de tous les k-jets des germes des applications de M dans N et on le note

$$J^k(M, N) = \text{espace des } k\text{-jets des applications de } M \text{ dans } N.$$

E x e m p l e. $J^1(\mathbb{R}, \mathbb{R})$ est un espace à trois dimensions; on peut le munir des coordonnées (x, y, p) (cf. § 3).

L'ensemble $J^k(M, N)$ possède une structure naturelle de variété différentiable. En effet, choisissons des systèmes de coordonnées au voisinage d'un point de M et au voisinage de l'image de ce point dans N par une application f. Alors les k-jets de l'application f et tous les jets voisins sont donnés par les coordonnées du point source et des collections de coefficients du polynôme de Taylor de l'application en ce point. Nous avons ainsi construit une carte de la variété des jets $J^k(M, N)$ au voisinage de son point qui est k-jet de l'application f.

Il est facile de calculer la dimension de la variété des jets. Par exemple,

$$J^0(M, N) = M \times N, \ \dim J^0(M, N) = \dim M + \dim N,$$
$$\dim J^1(M, N) = \dim M + \dim N + \dim M \dim N.$$

On a l'application naturelle

$$J^{k+1}(M, N) \to J^k(M, N)$$

(un $(k+1)$-jet définit un k-jet, puisque une $(k+1)$-tangence entraîne une k-tangence). Cette application différentiable est une fibration. On obtient la chaîne de fibrations

$$\ldots \to J^k \to J^{k-1} \to \ldots \to J^1 \to J^0 = M \times N.$$

La fibre de chacune de ces fibrations est difféomorphe à un espace vectoriel, mais ne possède pas (pour $k > 1$) de structure linéaire naturelle (« non-invariance des différentielles supérieures »).

Les variétés J^k sont une sorte d'approximations de dimensions finies de l'espace fonctionnel de dimension infinie des applications différentiables de M dans N.

D. Groupes des jets des difféomorphismes locaux et espaces des jets de champs de vecteurs.

Considérons l'espace des jets $J^k(M, N)$. Cette variété contient la sous-variété des k-jets des difféomorphismes. Cette sous-variété n'est pas un groupe, car les jets ne peuvent être multipliés que si le but d'un jet est la source d'un autre.

Fixons un point de M et considérons tous les germes des difféomorphismes de M laissant ce point fixe. Leurs k-jets forment un groupe.

D é f i n i t i o n. Le groupe des k-jets des germes des difféomorphismes de M laissant un point x invariant est appelé *groupe des k-jets des difféomorphismes locaux de la variété M au point x* et se note $J_x^k(M)$.

E x e m p l e. Le groupe des 1-jets des difféomorphismes locaux est isomorphe au groupe linéaire $J_x^1(M^m) = \mathrm{GL}(\mathbb{R}^m)$.

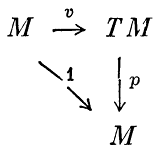
Pour $k > 1$ on obtient un groupe de Lie plus complexe. Vu qu'un k-jet définit un $(k-1)$-jet, on obtient la chaîne d'applications

$$\to J_x^k(M) \to \ldots \to J_x^1(M) = \mathrm{GL}(\mathbb{R}^m).$$

Il est aisé de voir que *ces applications* (les applications consistant à oublier les termes de degré k dans la série de Taylor) *sont des homomorphismes, et leurs noyaux, des groupes commutatifs.* Soit, par exemple, $m = 1$. On a:

◀ Si $f(x) = x + ax^k \pmod{(x^{k+1})}$ et $g(x) = x + bx^k \pmod{(x^{k+1})}$ alors $(f \circ g)(x) = x + ax^k + bx^k \pmod{x^{k+1}}$. ▶

Un *champ de vecteurs* sur une variété M est une section du fibré tangent p: $TM \to M$, c'est-à-dire une application différentiable v: $M \to TM$, telle que soit commutatif le diagramme



Les définitions des germes, jets et espaces des jets des champs de vecteurs répètent les définitions précédentes.

Le groupe des difféomorphismes de la variété M agit sur l'espace de tous les champs de vecteurs sur M, ainsi que sur les espaces des k-jets des champs de vecteurs sur M.

Le groupe des k-jets des difféomorphismes locaux de la variété M en un point agit sur l'espace des $(k-1)$-jets des champs de vecteurs sur M en ce point; cette action est linéaire.

E x e m p l e. Soit $y = a_1x + a_2x^2 + \ldots$ un 2-jet d'un difféomorphisme local au point 0. L'image d'un 1-jet du champ $v(x) = v_0 + v_1x + \ldots$ est donné par la formule $w(x) = w_0 + w_1x + \ldots$, où $w_0 = a_1v_0$, $w_1 = a_1v_1a_1^{-1} + 2a_2a_1^{-1}v_0$.

◀ On obtient cette formule en écrivant l'équation $\dot{x} = v(x)$ dans les coordonnées y. ▶

E. Théorème faible de transversalité.

La démonstration de la réductibilité au cas générique peut souvent être remplacée par une référence à certains théorèmes standards (et évidents) de transversalité. Plus bas on formule et on ébauche les démonstrations des théorèmes de transversalité les plus courants. On se refère aux théorèmes de transversalité essentiellement par souci d'économie de place: dans chaque cas concret la proposition correspondante se démontre directement sans peine.

D é f i n i t i o n. Deux sous-espaces vectoriels X et Y d'un espace vectoriel L sont *transversaux* si leur somme est l'espace L:

$$L = X + Y.$$

Par exemple, deux plans d'un espace à trois dimensions se coupant sous un angle non nul sont transversaux, deux droites, non.

Supposons que A et B sont des variétés différentiables et C une sous-variété différentiable de B (ici et dans la suite on admet que les variétés sont sans bord).

D é f i n i t i o n. On dit qu'une application $f: A \to B$ est transversale à C en un point a de A si ou bien $f(a) \notin C$, ou bien le plan tangent à C au point $f(a)$ et l'image du plan tangent à A au point a sont transversaux (fig. 105):

$$f_* T_a A + T_{f(a)} C = T_{f(a)} B.$$

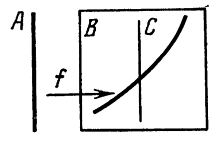


Fig. 105

D é f i n i t i o n. Une application $f: A \to B$ est *transversale* à C si elle lui est transversale en tout point de la variété source.

Par exemple, le plongement d'une droite dans un espace à trois dimensions est transversale à une autre droite dans cet espace si et seulement si ces deux droites ne sont pas concourantes.

R e m a r q u e. L'application d'une droite sur un plan peut ne pas être transversale à une autre droite de ce plan même si l'image de cette application est une normale à cette autre droite (l'image d'un espace tangent et d'un espace tangent à l'image ne représentent pas la même chose).

A noter encore que si $f: A \to B$ est transversale à C, la contre-image de C dans A est une sous-variété différentiable et sa codimension dans A est égale à celle de C dans B.

On rencontre souvent le cas où C n'est pas une sous-variété différentiable, mais une sous-variété présentant des singularités.

D é f i n i t i o n. On appelle *sous-variété stratifiée* d'une variété différentiable une réunion finie de variétés différentiables deux à deux disjointes (strates) vérifiant la condition suivante: l'adhérence de chaque strate est constituée de cette dernière et d'une réunion finie de strates de dimension plus petite.

Une application est *transversale à une sous-variété stratifiée* si elle est transversale à chaque strate.

E x e m p l e. Soit C réunion de deux plans sécants dans un espace à trois dimensions: la stratification est la partition en la droite d'intersection et quatre demi-plans. La transversalité à C équivaut à la transversalité à chacun de ces plans et à la droite d'intersection. Par exemple, une courbe transversale à la variété stratifiée C ne coupe pas la droite singulière de C.

Théorème. *Soient A variété compacte, C sous-variété compacte d'une variété B. Alors les applications $f: A \to B$ transversales à C forment un ensemble ouvert partout dense dans l'espace de toutes les applications assez lisses $A \to B$* (le voisinage d'une application f est défini comme le voisinage de la collection des fonctions définissant f et de leurs dérivées jusqu'à un ordre r assez grand).

Ce théorème s'appelle théorème faible de transversalité. Il affirme que par une petite déformation on peut transformer une application non transversale à une sous-variété fixe en transversale (fig. 106). Si la transversalité a lieu elle est préservée par de petites déformations.

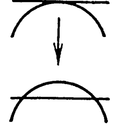


Fig. 106

◀ Voyons le cas particulier où B est un espace vectoriel, $B = \mathbb{R}^n$ et $C = \mathbb{R}^{n-k}$.

Ecrivons B sous la somme $B = C + D$ de deux sous-espaces de dimension complémentaire, $C = \mathbb{R}^{n-k}$ et $D = \mathbb{R}^k$. Projetons B sur D le long de C; désignons cette projection par π. Considérons l'application $\pi \circ f: A \to D$.

Le point 0 est valeur critique de cette application si et seulement si l'application $f: A \to B$ n'est pas transversale à la sous-variété $C \subset B$. D'après le lemme de Sard (§ 10) presque tous les points de D ne sont pas valeurs critiques de l'application $\pi \circ f$. Supposons que ε est un point de D valeur non critique pour $\pi \circ f$. Construisons l'application $f_\varepsilon: A \to B$ en posant $f_\varepsilon(a) = f(a) - \varepsilon$. Alors l'application f_ε est transversale à C. Comme ε peut être pris arbitrairement petit on déduit que les applications transversales forment un ensemble partout dense dans le cas particulier étudié. Le théorème des fonctions implicites nous apprend que cet ensemble est ouvert. Le cas général se ramène sans peine au cas étudié. ▶

Remarque. Si C n'est pas compacte il faut en général remplacer « ouvert » par « intersection d'un ensemble dénombrable d'ouverts ».

Exemple 1. B est un tore, C son hélice, A un cercle.

Exemple 2. B est un plan, A un cercle plongé dans le plan, C une tangente à ce cercle (sans le point de tangence). Le plongement est transversal à C, mais il existe des applications aussi proches que l'on veut non transversales à C.

Pour que les applications d'un compact A dans B, transversales à C forment un ensemble *ouvert* partout dense, il suffit, au lieu de la compacité de C d'exiger qu'en chaque point de B il existe un voisinage tel que le couple composé d'un voisinage et de son intersection avec C soit difféomorphe au couple $(\mathbb{R}^b, \mathbb{R}^c)$ ou bien au couple $(\mathbb{R}^b$, l'ensemble vide).

Si A n'est pas compacte, alors il est commode de munir l'espace des applications d'« une topologie fine ». Dans cette topologie, le voisinage de l'application $f: A \to B$ est définie de la manière suivante: fixons un ensemble ouvert G dans l'espace des jets $J^k(A, B)$ pour un k. L'ensemble des C^∞-applications $f: A \to B$ dont les k-jets en chaque point appartiennent à G, est ouvert pour la topologie fine. De tels ensembles ouverts vides sont pris pour base des voisinages définissant la topologie fine dans l'espace des C^∞-applications.

Donc, pour la topologie fine deux applications sont voisines si elles se rapprochent aussi vite que l'on veut (avec un nombre quelconque de dérivés)

« à l'infini »; en particulier, le graphe d'une application assez proche de f est situé dans un voisinage du graphe de l'application f qui s'amincit arbitrairement vite « à l'infini».

D'où il résulte que la convergence d'une suite pour la topologie fine entraîne la coïncidence totale, en dehors d'un compact, de tous les termes, de la suite à partir de l'un entre eux. Néanmoins, tout voisinage de cette application dans la topologie fine contient des applications qui ne sont nulle part confondues avec l'application donnée.

Si les applications forment un espace ouvert et partout dense pour la topologie fine, le théorème de transversalité est valable pour des variétés A non compactes (pour que cet espace soit ouvert il faut toutefois que C est compacte ou vérifie la condition formulée plus haut).

Le théorème de transversalité se généralise de façon évidente à la sous-variété stratifiée C. Cependant dans ce cas le théorème affirme que les applications transversales forment non pas un ouvert partout dense, mais seulement une intersection partout dense d'un nombre dénombrables d'ensembles ouverts.

Pour que les applications transversales à une variété stratifiée analytique forment un ensemble ouvert partout dense, il suffit que la stratification vérifie la condition subsidiaire suivante: toute immersion transversale à une strate est transversale à toutes les strates adjacentes de dimension supérieure dans un voisinage de cette strate.

Exemple 1. Soit C une réunion finie de plans dans un espace vectoriel, stratifié naturellement (par exemple, un couple de plans sécants de $\mathbb{R}^3$). La transversalité à $\mathbb{R}^k$ entraîne la transversalité à l'espace ambiant $\mathbb{R}^l$, donc notre condition est réalisée.

Exemple 2. Soit C le cône $x^2 = y^2 + z^2$ dans $\mathbb{R}^3$, la stratification est la décomposition en le point 0 et deux nappes. On vérifie aisément que notre condition est réalisée.

Exemple 3. Soit C un parapluie de Whitney défini par l'équation *) $y^2 = zx^2$ dans $\mathbb{R}^3$ (fig. 107). [La partie $z \geqslant 0$ de cette variété stratifiée est l'image de l'application $\varphi: \mathbb{R}^2 \to \mathbb{R}^3$ définie par les formules $x = u$, $z = v^2$, $y = uv$. Whitney a prouvé que 1) par une petite déformation l'application φ préserve les singularités du même type (aux difféomorphismes de $\mathbb{R}^2$ et $\mathbb{R}^3$ près); 2) que c'est l'unique singularité des applications de variétés à deux dimensions dans des variétés à trois dimensions, préservée par de petites déformations (à l'exception des lignes de self-intersection), toutes les autres singularités se désagrègent par une petite déformation en singularité de ce type.] La transversalité à la droite singulière $x = y = 0$ n'entraîne pas la transversalité à la variété des points réguliers de la surface proches de cette droite (le plan $z = 0$ est transversal à la droite et non transversal à la surface).

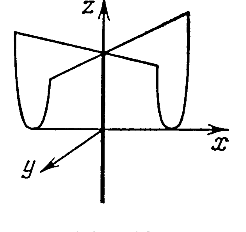


Fig. 107

Si la condition portant sur la stratification C: transversalité à une strate $\Rightarrow$ transversalité à toutes les strates de dimension supé-

*) Le parapluie comprend la canne représentée par une ligne grasse sur la figure 107.

rieure, est réalisée, alors la transversalité à toute la stratification est réalisée comme suit.

1) Les strates de dimension minimale ne présentent pas de singularités et sont justiciables du théorème ordinaire. 2) Au voisinage des strates de dimension minimale la transversalité est réalisée sur toutes les strates. 3) Eliminons de la variété ambiante l'adhérence des voisinages des strates de dimension minimale et passons aux strates de la dimension suivante.

Exemple. Soient B espace des opérateurs linéaires $b: \mathbb{R}^m \to \mathbb{R}^n$, C ensemble des opérateurs de rang non maximal. Les opérateurs de rang r forment une sous-variété différentiable dont la codimension est égale à $(m - r)(n - r)$ dans l'espace B. La décomposition en variétés d'opérateurs de rang divers définit une stratification sur C.

L'application $f: A \to B$ est une famille d'opérateurs linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^n$, dépendant différentiablement d'un point de la variété A comme d'un paramètre. La variété A est dite base de la famille. Le théorème de transversalité entraîne immédiatement le

Corollaire. *Dans l'espace des familles différentiables de matrices d'ordre $m \times n$, les familles transversales à la variété stratifiée C de matrices de rang non maximal forment un ensemble partout dense.*

En particulier, les valeurs du paramètre auquel correspondent les matrices de rang r forment généralement (pour les familles d'une intersection partout dense d'ensembles ouverts dans l'espace des familles) une sous-variété différentiable de codimension $(m - r) \times \times (n - r)$ dans la base de la famille.

Par exemple, dans une famille générique à cinq paramètres de matrices d'ordre 2×3 le rang baisse jusqu'à 1 sur une sous-variété différentiable tridimensionnelle de l'espace des paramètres et n'est égal à 0 en aucun point de cet espace ; s'il en va autrement pour cette famille, par une déformation aussi petite que l'on veut de la famille, on peut rendre cette famille générique.

F. Théorème de transversalité de Thom.

Le théorème de transversalité de Thom est une généralisation du théorème faible de transversalité dans lequel le rôle de la sous-variété C incombe à une sous-variété de l'espace des jets.

Chaque application différentiable $f: M \to N$ est rattachée à sa k-jet-extension $\hat{f}: M \to J^k(M, N)$, $\hat{f}(x) = j_x^k(f)$ (à un point x de M est associé le k-jet de l'application f en x).

Théorème. *Soit C une sous-variété de l'espace des jets $J^k(M, N)$. Alors l'ensemble des applications $f: M \to N$ dont les k-jets-extensions sont transversales à C est une intersection dénombrable partout dense d'ensembles ouverts dans l'espace de toutes les applications différentiables de M dans N.*

Ce théorème signifie que par une petite déformation on peut rendre une application différentiable générique non seulement par rapport à une sous-variété différentiable quelconque dans l'espace-image et encore par rapport à une condition quelconque portant sur les dérivées d'un ordre fini quelconque.

Remarque. Le théorème faible de transversalité se déduit à partir du théorème formulé pour $k = 0$. Le théorème fort ne découle pas immédiatement du théorème faible pour la raison suivante. On aurait pu appliquer le théorème faible à l'application $\hat{f}: M \to J^k$ et obtenir une application proche de $\hat{f}$, transversale à C. Cependant cette application proche ne sera pas généralement une k-jet-extension d'aucune application différentiable de M dans N.

Le théorème de transversalité de Thom affirme qu'on peut choisir une déformation transversalisante dans une classe plus étroite de déformations: il suffit de se borner à une déformation d'une k-jet-extension dans l'espace des k-jets-extensions et non dans l'espace de toutes les sections $M \to J^k$. Donc, le théorème signifie que les conditions d'intégrabilité (dont la réalisation établit une distinction entre les k-jets-extensions des applications de M dans N et les sections arbitraires $M \to J^k$) ne sont pas une entrave à la transversalité.

◀ Le principe de la démonstration consiste en une réduction au lemme de Sard comme pour le théorème faible de transversalité. La seule différence est que la déformation transversalisante est cherchée non pas dans la classe des applications $f_\varepsilon = f - \varepsilon$, mais dans une classe plus large de déformations polynomiales $f_\varepsilon = f + \varepsilon_1 e_1 + \ldots + \varepsilon_s e_s$, où e_i sont les monômes vectoriels de puissance $\leqslant k$.

Lemme 1. *Soit l'application différentiable $F: A \times E \to B$ du produit direct des variétés A et E dans la variété différentiable B.*

On envisagera F comme une famille d'applications F_ε des variétés A dans B, dépendant d'un point ε de la variété E comme d'un paramètre. Si l'application F est transversale à une sous-variété C de la variété B alors presque chaque membre $F_\varepsilon: A \to B$ correspondant à F de la famille est transversal à C.

◀ Considérons $F^{-1}(C)$. Le théorème des fonctions implicites nous dit que c'est une sous-variété différentiable dans $A \times E$. Considérons la projection de cette sous-variété dans E le long de A. D'après le lemme de Sard presque toutes les valeurs ne sont pas critiques. Soit ε valeur non critique, alors $F_\varepsilon: A \to B$ est transversale à C (car F est transversale à C et $A \times \varepsilon$ est transversale à $F^{-1}(C)$). ▶

Lemme 2. *Soit f une application différentiable de $\mathbb{R}^m$ dans $\mathbb{R}^n$. Fixons dans $\mathbb{R}^m$ et $\mathbb{R}^n$ des systèmes de coordonnées et considérons une application différentiable du produit direct de l'espace $\mathbb{R}^m$ par l'espace $\mathbb{R}^s$ dans l'espace des k-jets des applications $J^k(M, N)$, définie par*

$$(x, \varepsilon) \mapsto (j_x^k f_\varepsilon),$$

où $f_\varepsilon = f + \varepsilon_1 e_1 + \ldots + \varepsilon_s e_s$ *($e_1, \ldots, e_s$ représentent tous les produits possibles des monômes de degré $\leqslant k$ des coordonnées d'un point x de $\mathbb{R}^m$ sur les vecteurs de base dans $\mathbb{R}^n$).*

On affirme que l'application construite ne possède pas de valeurs critiques (et, par conséquent, est transversale à toute sous-variété de l'espace des k-jets).

◀ Les coordonnées de l'espace J^k sont les coordonnées du point x de $\mathbb{R}^m$ et les coefficients de Taylor d'un jet dans ce point jusqu'à l'ordre k compris. Par un choix convenable des coefficients $\varepsilon_1, \ldots, \varepsilon_s$ le polynôme vectoriel $\varepsilon_1 e_1 + \ldots + \varepsilon_s e_s$ aura en un point quelconque donné à priori une collection quelconque donnée à priori de coefficients de Taylor jusqu'aux membres d'ordre k inclus. D'où le lemme. ▶

Soit C une sous-variété différentiable dans $B = J^k(\mathbb{R}^m, \mathbb{R}^n)$. Appliquons à l'application du lemme 2 le lemme 1 (dans lequel $A = \mathbb{R}^m$, $E = \mathbb{R}^s$; $F(x, \varepsilon) = j_x^k f_\varepsilon$). D'après le lemme 1 l'application $F_\varepsilon = F(\cdot, \varepsilon)$ est transversale à C pour presque tous les ε. En choisissant ε assez petit on peut obtenir une application f_ε: $\mathbb{R}^m \to \mathbb{R}^n$ aussi proche que l'on veut de f (dans un sous-ensemble fini quelconque de $\mathbb{R}^m$) et dont la k-jet-extension est transversale à C. Le passage de cette construction locale à la construction globale (la substitution de $\mathbb{R}^m$, $\mathbb{R}^n$ à M, N) ne présente pas de difficultés. ▶

G. Exemple : Désagrégation des points singuliers dégénérés d'un champ de vecteurs.

A titre d'application des théorèmes de transversalité nous allons voir quels points singuliers est susceptible de posséder un champ de vecteur générique.

Définition. Un point singulier x d'un champ de vecteurs v est *non dégénéré* si l'opérateur de la partie linéaire du champ est non dégénéré en ce point.

Les théorèmes de transversalité entraînent le

Corollaire. *Dans l'espace fonctionnel des champs de vecteurs différentiables sur une variété compacte, les champs dont les points singuliers ne sont pas dégénérés (et par conséquent, sont isolés) forment un ensemble ouvert partout dense.*

◀ Les points singuliers sont les antécédants d'une variété différentielle (la section nulle) dans l'espace des 0-jets des champs de vecteurs. La non-dégénérescence d'un point singulier est la transversalité de la 0-jet-extension du champ à cette variété. ▶

Donc, un point singulier dégénéré se désagrège en points non dégénérés par une déformation aussi petite que l'on veut du champ.

Exemple. Soit un point singulier du type « col-nœud »:

$$\dot{x} = x^2, \quad \dot{y} = -y.$$

Par la déformation $\dot{x} = x^2 - \varepsilon$, $\dot{y} = -y$ le col-nœud se désagrège en deux points singuliers: un col et un nœud.

Il se pose la question de savoir en combien de points singuliers se désagrège un point singulier dégénéré donné par de petites déformations. Comme c'est souvent le cas (par exemple dans la théorie des équations algébriques), ce problème se résout le plus naturellement dans le domaine complexe.

Définition. On appelle *multiplicité* d'un point singulier dégénéré le nombre de points singuliers non dégénérés (complexes) en lesquels se désagrège ce point par une petite déformation.

Remarque. En toute rigueur la multiplicité se définit de la manière suivante: 1) on fixe un voisinage assez petit du point singulier dans l'espace complexe; 2) en fonction du voisinage retenu on choisit une petite déformation; 3) on compte le nombre de points singuliers du champ déformé au voisinage du point donné.

Plus bas on indiquera une formule pour la multiplicité d'un point singulier en termes de diagrammes de Newton (A. Kouchnirenko, D. Bernstein, A. Hovanski).

Soit $f = \sum f_m x^m$ une série numérique formelle entière en les variables $x_1, \ldots, x_n$ ($x^m = x_1^{m_1} \ldots x_n^{m_n}$). Considérons l'octante du réseau des points entiers m à coordonnées non négatives m_k.

Notons cet octante $\mathbb{Z}_+^n$.

Définition. On appelle *support* de la série f l'ensemble des points $m \in \mathbb{Z}_+^n$ pour lesquels $f_m \neq 0$. On note

$$\operatorname{supp} f = \{m \in \mathbb{Z}_+^n : f_m \neq 0\}.$$

Définition. On appelle *polyèdre de Newton* de la série f l'enveloppe convexe dans l'octante $\mathbb{R}_+^n$ de la réunion des octantes parallèles à $\mathbb{Z}_+^n$ et dont les sommets sont les points du support de l'espace vectoriel réel. On note:

$$\Gamma_f = \text{enveloppe convexe de la réunion } m + \mathbb{Z}_+^n,\ m \in \operatorname{supp} f.$$

Un polyèdre de Newton est *commode* s'il coupe tous les axes de coordonnées.

Théorème. *Soient donnés n polyèdres de Newton commodes $\Gamma_1, \ldots, \Gamma_n$.*

Considérons le champ de vecteurs $v_1\partial/\partial x_1 + \ldots + v_n\partial/\partial x_n$, où les polyèdres de Newton pour les composantes $v_1, \ldots, v_n$ sont $\Gamma_1, \ldots, \Gamma_n$. Alors la multiplicité μ du point singulier 0 du champ de vecteurs n'est pas inférieure au nombre de Newton ν $(\Gamma_1, \ldots, \Gamma_n)$ qui sera défini plus bas; elle coïncide avec lui pour tous les champs, à l'exception d'une hypersurface dans l'espace des champs ayant pour polyèdres de Newton les polyèdres donnés.

R e m a r q u e. La condition de commodité des polyèdres n'est pas restrictive, car on peut démontrer qu'elle peut être réalisée par l'adjonction de termes de puissances aussi grandes que l'on veut ne modifiant pas la multiplicité (pourvu qu'elle soit finie).

Pour définir le nombre de Newton d'un système de polyèdres commodes on aura besoin de la notion de volume mixte.

Soit Γ un polyèdre commode de Newton. Par *volume* $V(\Gamma)$ on entendra le volume du domaine (non convexe) compris entre le zéro et la frontière du polyèdre Γ dans l'octante positive $\mathbb{R}^n_+$.

Soient Γ_1 et Γ_2 deux polyèdres commodes de Newton. On appelle somme arithmétique la somme $\Gamma_1 + \Gamma_2$, c'est-à-dire l'ensemble de toutes les sommes possibles de vecteurs de Γ_1 et de Γ_2. Cette somme est également un polyèdre commode de Newton.

Donc, les polyèdres commodes de Newton forment un semi-groupe commutatif. A partir de ce semi-groupe, on peut construire un groupe (appelé groupe de Grothendieck) par la procédure habituelle: la différence formelle de deux polyèdres de Newton est un élément du groupe et, de plus, par définition, $\Gamma_1 - \Gamma_2 = \Gamma_3 - \Gamma_4$ si et seulement si $\Gamma_1 + \Gamma_4 = \Gamma_2 + \Gamma_3$.

Le groupe construit forme également un espace vectoriel sur le corps des réels: si λ est un nombre positif alors $\lambda\Gamma$ désigne le polyèdre homothétique de Γ par une homothétie de centre 0 et de coefficient λ. Le volume $V(\Gamma)$ est prolongeable de façon unique à l'espace vectoriel comme forme de degré m (la démonstration de ce fait, qui n'est pas complètement évident, est laissé à titre d'exercice au lecteur curieux).

Chaque forme de degré m se représente d'une manière unique par les valeurs d'une forme n-linéaire symétrique lorsque les arguments coïncident: par exemple,

$$a^2 = ab \mid a = b, \quad ab = ((a + b)^2 - a^2 - b^2)/2.$$

D é f i n i t i o n. On appelle *volume mixte* (de Minkowski) d'un système de polyèdres $(\Gamma_1, \ldots, \Gamma_n)$ la valeur prise sur $(\Gamma_1, \ldots, \Gamma_n)$ par l'unique forme n-linéaire symétrique qui coïncide avec le volume $V(\Gamma)$ pour $\Gamma_1 = \ldots = \Gamma_n = \Gamma$. On note: $V(\Gamma_1, \ldots, \Gamma_n)$.

E x e m p l e. Dans le plan $n = 2$ et le volume mixte du couple (Γ_1, Γ_2) est $V(\Gamma_1, \Gamma_2) = [V(\Gamma_1 + \Gamma_2) - V(\Gamma_1) - V(\Gamma_2)]/2$.

D é f i n i t i o n. Le *nombre de Newton* $\nu(\Gamma_1, \ldots, \Gamma_n)$ se définit comme suit

$$\nu(\Gamma_1, \ldots, \Gamma_n) = n!V(\Gamma_1, \ldots, \Gamma_n).$$

E x e m p l e. Supposons que dans le plan, Γ_1 et Γ_2 sont limitées par les droites coupant les axes des coordonnées aux points (a_1, b_1) pour Γ_1, et (a_2, b_2) pour Γ_2. Alors $\nu(\Gamma_1, \Gamma_2) = \min(a_1b_2, a_2b_1)$. Donc, la multiplicité d'un point singulier est génériquement égale à

$$\mu = \min(a_1b_2, a_2b_1).$$

§ 30. Matrices dépendant de paramètres et singularités du diagramme du décrément

Pour préparer l'étude des bifurcations des points singuliers de champs de vecteur nous considérons ici le problème de la forme normale de familles d'endomorphismes d'un espace vectoriel.

A. Problème de la forme normale de matrices dépendant de paramètres.

La réduction d'une matrice à la forme normale de Jordan n'est pas une opération stable si parmi les valeurs propres existent des valeurs multiples. En effet, par une variation aussi petite que l'on veut de la matrice on peut modifier totalement la forme de Jordan si les valeurs propres sont multiples. Donc, si une matrice n'est connue qu'approximativement, sa réduction à la forme normale de Jordan est pratiquement impossible en cas de valeurs propres multiples. Cette réduction n'est pas indispensable puisque la matrice générique ne possède pas de valeurs propres multiples.

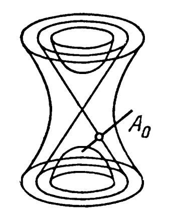


Fig. 108

Les valeurs propres multiples sont inévitables par une petite déformation dans le cas où l'on s'intéresse non pas à une matrice isolée, mais à toute une famille des matrices dépendant des paramètres. Dans ce cas même si nous pouvons réduire à la forme normale de Jordan chaque matrice de la famille, cette forme normale et la transformation qui ramène à cette forme dépendront en général discontinûment du paramètre.

Il se pose donc la question de savoir à quelle forme simple on peut ramener une famille de matrices dépendant différentiablement (pour fixer les idées holomorphement) des paramètres par des changements de coordonnées dépendant différentiablement (holomorphement) des paramètres.

Traitons l'ensemble de toutes les matrices complexes carrées d'ordre n comme un espace vectoriel de dimension n^2. La relation de similitude des matrices décompose l'espace $\mathbb{C}^{n^2}$ en variétés (orbites du groupe linéaire $GL(n, \mathbb{C})$: deux matrices sont situées sur une même orbite si leurs valeurs propres et les dimensions de leurs blocs de Jordan sont les mêmes. Cette partition est continue en raison de l'existence des valeurs propres. A titre de modèle grossier on peut s'imaginer la partition de l'espace $\mathbb{R}^3$ en strates des variétés $x^2 + y^2 - z^2 = C$ (fig. 108).

La famille de matrices est définie par une application de l'espace des paramètres de la famille dans l'espace des matrices $\mathbb{C}^{n^2}$. Il s'avère que parmi toutes les familles de matrices on peut exhiber des familles spéciales (auxquelles on peut réduire toutes les familles)

15—0743

telles que cette réduction soit réalisée par un changement de base dépendant différentiablement des paramètres (et un changement différentiable de paramètres). Ces familles sont dites déformations verselles (la définition exacte est donnée plus bas). Les déformations verselles à nombre minimal de paramètres sont dites miniverselles.

Donc, les déformations miniverselles sont les formes normales possédant le plus petit nombre de paramètres à la réduction desquelles on peut conserver la dépendance différentiable par rapport aux paramètres.

Exemple. Si toutes les valeurs propres d'une matrice diagonale sont distinctes, alors pour déformation miniverselle on peut prendre la famille de toutes les matrices diagonales (les paramètres sont des valeurs propres).

Nous indiquons plus bas les déformations miniverselles de matrices arbitraires.

B. Déformations verselles.

Définition. On appellera *famille de matrices* une application holomorphe $A: \Lambda \to \mathbb{C}^{n^2}$, où Λ est un voisinage de l'origine des coordonnées dans un espace des paramètres $\mathbb{C}^l$. On appellera *déformation* de la matrice A (0) le germe de la famille A au point 0.

Une déformation A' de la matrice A (0) est *équivalente* à la déformation A si existe une déformation de l'unité C, telle que

$$A'(\lambda) = C(\lambda) A(\lambda) (C(\lambda))^{-1}.$$

Soit φ: $(M, 0) \to (\Lambda, 0)$ application holomorphe ($M \subset \mathbb{C}^m$, $\Lambda \subset \mathbb{C}^l$).

Définition. On appelle famille *induite par A par l'application* φ la famille φ^*A :

$$(\varphi^*A)(\mu) = A(\varphi(\mu)), \quad \mu \in M.$$

La déformation induite φ^*A de la matrice A (0) est définie par la même formule.

Définition. Une déformation A d'une matrice A_0 est dite *verselle* si toute déformation A' de la matrice A_0 est équivalente à une déformation induite par A. Une déformation verselle est *universelle* si cette application induite φ est déterminée de façon unique par la déformation A'. Une déformation verselle est dite *miniverselle* si la dimension de l'espace des paramètres est la plus petite possible pour toute déformation verselle.

Exemple. Une famille de matrices diagonales d'éléments diagonaux $(\alpha_i + \lambda_i)$, où tous les α_i sont distincts et λ_i les paramètres de la déformation, est une déformation verselle, universelle et miniverselle de la matrice (α_i).

La famille de toutes les matrices de $\mathbb{C}^{n^2}$ définit une déformation verselle n^2-paramétrique de l'une quelconque de ses matrices. Cepen-

dant cette déformation n'est généralement ni universelle ni miniverselle.

La dimension d'une déformation miniverselle d'une matrice quelconque est donnée par le théorème suivant. Soient α_i les valeurs propres de la matrice A_0 et $n_1(\alpha_i) \geqslant n_2(\alpha_i) \geqslant \ldots$ les dimensions des blocs de Jordan de valeurs propres α_i, classées par ordre de décroissance.

Théorème 1. *Le nombre minimal de paramètres d'une déformation verselle de la matrice A_0 est égal à*

$$\sum_i [n_1(\alpha_i) + 3n_2(\alpha_i) + 5n_3(\alpha_i) + \ldots].$$

Les déformations miniverselles peuvent être choisies de diverses façons. En particulier, les trois formes normales décrites dans le théorème suivant sont des déformations verselles d'une matrice réduite à une forme normale de Jordan triangulaire supérieure.

Théorème 2. *Soit A une famille dépendant holomorphement d'un paramètre $\lambda \in \mathbb{C}^l$ d'opérateurs linéaires de $\mathbb{C}^n$ dans $\mathbb{C}^n$ et supposons que pour une valeur λ_0 de λ l'opérateur $A(\lambda_0)$ possède les valeurs propres α_i et des blocs de Jordan d'ordre*

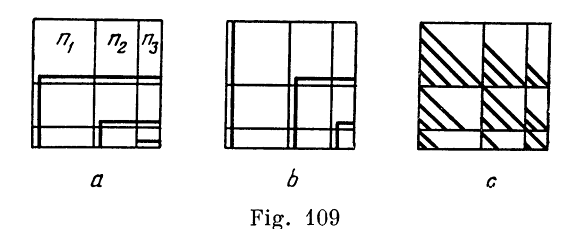


Fig. 109

$$n_1(\alpha_i) \geqslant n_2(\alpha_i) \geqslant \ldots$$

Il existe alors une base dans $\mathbb{C}^n$ dépendant holomorphement du paramètre λ, variable dans un voisinage du point λ_0, et telle que la matrice de l'opérateur $A(\lambda)$ y prenne la forme diagonale

$$A_0 + B(\lambda),$$

où A_0 est une matrice de Jordan triangulaire supérieure de l'opérateur $A(\lambda_0)$, et $B(\lambda)$ une matrice diagonale de Jordan dont les blocs correspondent aux valeurs propres de la matrice A_0.

Le bloc B_i correspondant à la valeur propre α_i est composé de zéros à l'exception des endroits signalés sur la figure 109, où l'on trouve des fonctions holomorphes de λ.

La figure 109 représente trois formes normales. Dans les deux premières le nombre d'éléments non nuls de B est égal à $n_1(\alpha_i) + 3n_2(\alpha_i) + \ldots$; dans la troisième tous les éléments d'une oblique

sont égaux. On obtient les déformations miniverselles de la matrice A_0 si l'on suppose que les éléments repérés des matrices B_i sont des variables indépendantes. Dans les trois cas leur nombre est égal à $\sum [n_1(\alpha_i) + 3n_2(\alpha_i) + \ldots]$. L'avantage des deux premières formes normales est que le nombre des éléments non nuls de la matrice est minimal. L'avantage de la troisième forme est l'orthogonalité de la déformation verselle à l'orbite correspondante (au sens du produit scalaire des matrices termes à termes).

C. Démonstration de la versalité.

◀ Soit $A : \Lambda \to \mathbb{C}^{n^2}$ une déformation de la matrice $A_0 = A(0)$ de paramètre $\lambda \in \Lambda$, telle que l'application A soit transversale à l'orbite C de la matrice A_0 par l'action du groupe de changements linéaires des coordonnées. Supposons que le nombre de paramètres de la déformation est minimal (c'est-à-dire égal à la codimension de l'orbite dans l'espace $\mathbb{C}^{n^2}$ de toutes les matrices). Une telle déformation est dite minitransverselle.

Lemme 1. *Une déformation minitransverselle A est miniverselle.*

Pour démontrer ce lemme on aura besoin de la

Définition. On appelle *centralisateur* de la matrice u l'ensemble de toutes les matrices commutant avec u. On note

$$Z_u = \{v : [u, v] = 0\}, \quad [u, v] = uv - vu.$$

Le centralisateur de toute matrice d'ordre n est un sous-espace vectoriel de l'espace $\mathbb{C}^{n^2}$ de toutes les matrices d'ordre n.

Soit Z le centralisateur de la matrice A_0. Traçons à partir de l'unité e dans l'espace des matrices non dégénérées une surface différentiable transversale à $e + Z$, de même dimension que la codimension du centralisateur (c'est-à-dire ayant une dimension minimale). Désignons cette surface par P et considérons l'application

$$\Phi : P \times \Lambda \to \mathbb{C}^{n^2}, \quad \Phi(p, \lambda) = pA(\lambda)p^{-1}.$$

Lemme 2. *Au voisinage du point $(e, 0)$ l'application Φ est un difféomorphisme local sur $(\mathbb{C}^{n^2}, A_0)$.*

Pour démontrer le lemme 2 nous allons considérer l'application ψ du groupe des matrices non dégénérées $\mathbb{C}^{n^2}$ de toutes les matrices, définie par la formule $\psi(b) = bA_0b^{-1}$.

1°. *La dérivée de l'application ψ en l'unité est l'opérateur de commutation avec A_0:*

$$\psi_* : \mathbb{C}^{n^2} \to \mathbb{C}^{n^2}, \quad \psi_* u = [u, A_0].$$

$$◀ \ (e + \varepsilon u) A_0 (e + \varepsilon u)^{-1} = A_0 + \varepsilon [u, A_0] + \ldots \ ▶$$

Il suit de 1°.

2°. *La dimension du centralisateur de la matrice* A_0 *est égale à la codimension de l'orbite, quant à la dimension de la transversale au centralisateur, elle est égale à la dimension de l'orbite*:

$$\dim Z = \dim \Lambda, \quad \dim P = \dim C.$$

Introduisons dans l'espace $\mathbb{C}^{n^2}$ le produit scalaire hermitien $\langle A, B\rangle = \operatorname{tr}(AB^*)$, où B^* est la transposée conjuguée de B. Le carré scalaire correspondant est tout simplement la somme des carrés des modules des éléments de la matrice.

Lemme 3. *Un vecteur* B *de l'espace tangent à* $\mathbb{C}^{n^2}$ *au point* A_0 *est perpendiculaire à l'orbite de la matrice* A_0 *si et seulement si* $[B^*, A_0] = 0$.

◀ Les vecteurs tangents à l'orbite sont des matrices mises sous la forme $[X, A_0]$. L'orthogonalité de B à l'orbite signifie que l'on a $\langle [X, A_0], B\rangle = 0$ pour tout X. Autrement dit, pour tout X

$$0 = \operatorname{tr}([X, A_0], B^*) = \operatorname{tr}(XA_0B^* - A_0XB^*) =$$
$$= \operatorname{tr}([A_0, B^*], X) = \langle [A_0, B^*], X^*\rangle.$$

Cette condition équivaut à $[A_0, B^*] = 0$, puisque X est arbitraire.

Fig. 110

On a donc démontré le lemme: l'orthocomplément à une orbite de la matrice se déduit de son centralisateur par une transposition et une conjugaison. ▶

Il n'est pas difficile d'écrire les centralisateurs des matrices réduites à la forme normale de Jordan. Supposons d'abord que la matrice ne possède qu'une valeur propre et une série de blocs de Jordan d'ordre $n_1 \geqslant n_2 \geqslant \ldots$

Lemme 4. *Les matrices représentées sur la figure* 110 *et elles seules commutent avec la matrice* A_0.

Sur la figure 110 tout segment oblique représente une série de nombres égaux, les places vides étant supposées être des 0. Donc, le nombre de segments est égal à la dimension du centralisateur.

◀ Le lemme 4 se démontre par un calcul immédiat du commutateur (cf., par exemple, F. Gantmakher. Théorie des matrices, Editions Dunod). ▶

Le lemme 4 entraîne que la dimension du centralisateur de la matrice A_0 (qui est égale à la codimension de l'orbite et à la plus petite dimension de la déformation verselle) est donnée par la formule $d = n_1 + 3n_2 + 5n_3 + \ldots$

Si la matrice de Jordan A_0 possède plusieurs valeurs propres, partageons-la en blocs correspondant à ces valeurs propres. Alors les matrices commutant avec A_0 seront diagonales par blocs, et de plus à chaque valeur propre correspondra un bloc de la forme décrite sur la figure 110. Donc la formule de la dimension du centralisateur

(la codimension de l'orbite, la dimension de la déformation universelle) se déduit de la précédente par une sommation sur toutes les valeurs propres distinctes.

En effet, ψ_* est une application linéaire d'espaces de même dimension, donc la dimension du noyau est égal à la codimension de l'image.

Démonstration du lemme 2. ◀ La dérivée de Φ par rapport à p dans $(e, 0)$ est ψ_*, la dérivée par rapport à λ est A_*. D'après ce qu'on a démontré, ces opérateurs appliquent isomorphement les espaces tangents à P en e et à Λ en 0 dans des espaces transversaux de mêmes dimensions (l'espace tangent à l'orbite C en A_0 pour P et l'espace transversal à C pour Λ). Donc, la dérivée de Φ au point $(e, 0)$ est un isomorphisme d'espaces vectoriels de dimensions n^2. Le théorème de la fonction inverse nous apprend que Φ est un difféomorphisme local. ▶

Démonstration du lemme 1. ◀ On traitera p et λ comme les coordonnées du point $\Phi(p, \lambda)$. Soient $A'\colon (M, 0) \to (\mathbb{C}^{n^2}, A_0)$ une déformation quelconque de la matrice A_0, $\mu \in M$ le paramètre de la déformation. Définissons $\lambda = \varphi(\mu)$ par la formule $\varphi(\mu) = \lambda(A'(\mu))$ et posons $B(\mu) = p(A'(\mu))$. Alors

$$A'(\mu) = B(\mu)\, A(\varphi(\mu))\, B^{-1}(\mu),$$

d'où il suit que la déformation A est verselle.

Que la dimension de la base de cette déformation est minimale est un fait évident. ▶

Pour déformation transverselle de la matrice A_0 on peut prendre une famille de matrices de la forme $A_0 + B$, où la matrice B appartient à l'orthocomplément décrit plus haut à l'orbite de la matrice A_0. On obtient ainsi une déformation miniverselle de la matrice A_0.

Dans le cas où la matrice A_0 ne possède qu'une seule valeur propre, la matrice B est de la forme représentée sur la figure 109, *c*. Dans cette formule chaque segment oblique est constitué de nombres égaux ; le nombre de paramètres est égal au nombre de segments et est donné par la formule indiquée plus haut.

La matrice B possède de nombreux éléments non nuls. On peut construire les déformations miniverselles $A_0 + B$ dans lesquelles le nombre d'éléments non nuls de la matrice B est minimal (égal au nombre de paramètres). A cet effet, choisissons une base dans le centralisateur : associons à chaque segment oblique de la figure 109, *c* une matrice composée de zéros et d'unités de telle sorte que les unités soient situées sur ce segment oblique.

Les équations suivantes constituent un système d'équations indépendantes du plan tangent à une orbite : la somme des éléments correspondants de la matrice est égale à 0 (lemmes 3 et 4) pour chaque segment oblique de la figure 109. Donc, pour obtenir une famille $A_0 + B$, transversale à l'orbite, il suffit de prendre, pour famille

de matrices B, des matrices dans lesquelles chaque segment oblique de la figure 109, c est composé du paramètre indépendant et d'éléments nuls. Dans chaque segment oblique on peut choisir l'élément non nul dans un endroit quelconque. Par exemple le choix indiqué dans le théorème 2 du point B fait l'affaire. ▶

D. Exemples.

On désignera une matrice de Jordan triangulaire supérieure par le produit des déterminants de ses blocs. Par exemple, α^2 représente le bloc de Jordan d'ordre 2, $\alpha\alpha$, la matrice du deuxième ordre, multiple de l'unité.

La première forme normale du théorème du point B nous conduit aux déformations miniverselles suivantes:

a) La déformation verselle (et universelle) à deux paramètres du bloc de Jordan α^2 d'ordre 2:

$$\begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \lambda_1 & \lambda_2 \end{pmatrix}. \tag{1}$$

b) La déformation verselle (mais non universelle) à quatre paramètres de la matrice scalaire $\alpha\alpha$ d'ordre 2:

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix} + \begin{pmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4 \end{pmatrix}.$$

c) La déformation verselle et universelle à trois paramètres de bloc de Jordan α^3:

$$\begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix}.$$

d) La déformation verselle à cinq paramètres de la matrice

$$\begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_4 & 0 & \lambda_5 \end{pmatrix}.$$

Par exemple, toute famille holomorphe de matrices dans laquelle le bloc de Jordan α^2 correspond à la valeur 0 du paramètre, se ramène pour de valeurs proches du paramètre à la forme normale (1), où λ_1, λ_2 sont des fonctions holomorphes des paramètres.

Les formes normales construites permettent lorsqu'on étudie des problèmes rattachés au comportement d'opérateurs dépendant de paramètres de se limiter à des familles spéciales: les déformations miniverselles. L'un de ces problèmes est la construction de diagrammes des bifurcations.

E. Diagrammes des bifurcations.

On appellera diagramme des bifurcations d'une famille de matrices la partition de l'ensemble des paramètres Λ suivant les formes de Jordan des matrices. Une famille est une application $A : \Lambda \to \mathbb{C}^{n^2}$ de l'espace des paramètres dans l'espace des matrices, donc l'étude des diagrammes des bifurcations passe par celle de la partition de l'espace de toutes les matrices en matrices de Jordan de type différent. On groupera les matrices dont les dimensions des blocs de Jordan sont égales et qui ne diffèrent que par la grandeur des valeurs propres. Donc, la partition obtenue est une stratification finie de l'espace des matrices.

Chaque strate de cette stratification est définie par l'ensemble de collections $n_1(i) \geqslant n_2(i) \geqslant \ldots$ des dimensions des blocs de Jordan correspondant à ν valeurs propres distinctes ($1 \leqslant i \leqslant \nu$). La codimension c d'une telle strate dans l'espace $\mathbb{C}^{n^2}$ est inférieure à la codimension d de l'orbite correspondante d'un nombre égal au nombre de valeurs propres distinctes, c'est-à-dire à ν:

$$c = d - \nu = \sum_{i=1}^{\nu} [n_1(i) + 3n_2(i) + \ldots - 1].$$

On remarquera que la contribution des valeurs propres simples dans cette somme est nulle. En se servant du théorème faible de transversalité on obtient le

Théorème. *Les familles transversales à la stratification définie plus haut forment un ensemble partout dense dans l'espace des familles de matrices d'ordre n.*

Ce théorème et les formules des déformations verselles du point D permettent de décrire les diagrammes des bifurcations des familles génériques. En particulier, pour les familles à un nombre peu élevé de paramètres on est conduit aux résultats suivants.

1°. *Familles à un paramètre.* La relation $c = 1$ entraîne que la matrice ne possède qu'une valeur propre double à laquelle correspond un bloc de Jordan d'ordre 2. On désignera une telle strate par α^2.

Corollaire. *Une famille générique à un paramètre de matrices ne comporte comme membre général que des matrices à valeurs propres simples, et pour certaines valeurs isolées du paramètre, des matrices de type α^2 (avec un bloc de Jordan d'ordre 2). Si cette famille possède des matrices de structure de Jordan plus compliquée, on peut les éliminer par une déformation aussi petite que l'on veut de la famille.*

2°. *Familles à deux paramètres.* Il existe très exactement deux types de matrices de Jordan avec $c = 2$: α^3 (un bloc de Jordan d'ordre 3) et $\alpha^2\beta^2$ (deux blocs d'ordre 2 de valeurs propres distinctes).

Corollaire. *Le diagramme des bifurcations d'une famille générique à deux paramètres de matrices a la forme d'une courbe dont les seules singularités sont des points de rebroussement et des points de*

self-intersection (fig. 111). *Aux points de rebroussement correspondent des matrices de type α^3 avec un seul bloc de Jordan d'ordre* 3, *aux points de self-intersection, des matrices de type $\alpha^2\beta^2$ avec deux blocs de Jordan d'ordre* 2 *avec des valeurs propres distinctes, aux points de la courbe, des matrices à un bloc de Jordan d'ordre* 2, *aux points situés en dehors de la courbe, des matrices à valeurs propres simples.*

Si la famille renferme des matrices de types plus compliqués, ou bien si le diagramme des bifurcations possède des singularités plus compliquées, alors on peut les éliminer par une déformation aussi petite que l'on veut de cette famille.

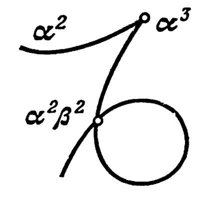


Fig. 111

3°. *Familles à trois paramètres.* Il existe 4 strates avec $c = 3$: $\alpha^2\beta^2\gamma^2$ (trois 2-blocs), $\alpha\alpha$ (deux 1-blocs de même valeur propre), $\alpha^2\beta^3$ (deux blocs d'ordre 2 et 3) et α^4 (un 4-bloc).

Donc, les points singuliers des diagrammes des bifurcations des familles génériques à trois paramètres sont de la forme représentée sur la figure 112. La singularité α^4 est dite queue d'aronde: cette surface est définie par l'équation $\Delta(a, b, c) = 0$, où Δ est le discriminant du polynôme $z^4 + az^2 + bz + c$. En toute rigueur, ce qui a été dit plus haut ne concerne que le cas complexe, puisque les surfaces représentées sur la figure 112 doivent être considérées comme des surfaces complexes.

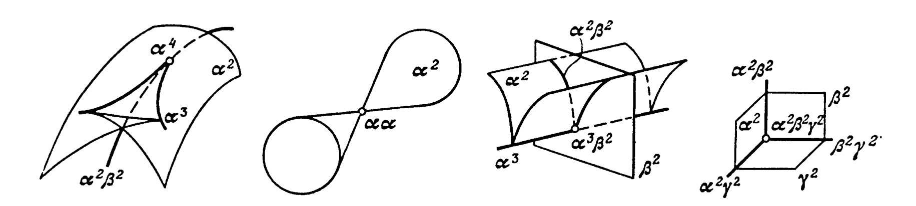


Fig. 112

Les déformations verselles de matrices réelles ont été construites par D. Galine (D. G a l i n e, Sur les matrices réelles dépendant de paramètres, UMN, 27, 1 (1972), 241-242, en russe). La construction s'effectue de la manière suivante. Supposons tout d'abord que l'opérateur réel dans $\mathbb{R}^{2n}$, pour lequel on cherche la déformation verselle, possède un couple unique de valeurs propres conjuguées complexes $x \pm iy$ $(y \neq 0)$ de blocs de Jordan de dimension $n_1 \geqslant n_2 \geqslant \ldots$, tels que $n_1 + n_2 + \ldots = n$. Alors la matrice de l'opérateur dans une base réelle de $\mathbb{R}^{2n}$ est de la même forme que la matrice du réélifié de l'opérateur complexe de Jordan $\hat{A}_0: \mathbb{C}^n \to \mathbb{C}^n$ d'unique valeur propre $x + iy$ et de blocs de Jordan de dimension $n_1 \geqslant n_2 \geqslant \ldots$,

c'est-à-dire de la forme

$$A_0 = \begin{pmatrix} X & -yE \\ yE & X \end{pmatrix}, \tag{2}$$

où X est une matrice de Jordan réelle triangulaire supérieure de valeur propre x et de blocs de dimensions $n_1 \geqslant n_2 \geqslant \ldots$, et E la matrice unité d'ordre n.

Il s'avère que *pour déformation verselle minimale de la matrice réelle A_0 on peut prendre la réélifiée d'une déformation verselle complexe minimale de la matrice complexe $\hat{A}_0$.*

Par exemple, pour déformation verselle minimale d'une matrice réelle d'ordre 4 de blocs de Jordan d'ordre 2 et de valeurs propres $x \pm iy$ on peut prendre une déformation à 4 paramètres obtenue par réélification de la déformation verselle complexe

$$\begin{pmatrix} z & 1 \\ 0 & z \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \lambda_1 & \lambda_2 \end{pmatrix},$$

c'est-à-dire la déformation de paramètres ρ_1, ρ_2, τ_1, τ_2:

$$\begin{pmatrix} x & 1 & -y & 0 \\ 0 & x & 0 & -y \\ y & 0 & x & 1 \\ 0 & y & 0 & x \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ \rho_1 & \rho_2 & -\tau_1 & -\tau_2 \\ 0 & 0 & 0 & 0 \\ \tau_1 & \tau_2 & \rho_1 & \rho_2 \end{pmatrix}, \quad z = x + iy, \quad \lambda_k = \rho_k + i\tau_k.$$

Chaque matrice réelle se ramène par une similitude sur le corps des nombres réels à une forme quasi diagonale dans laquelle à toute valeur propre réelle est associée une matrice de Jordan réelle et à chaque couple de valeurs propres conjuguées complexes, un bloc de la forme (2).

On obtient une déformation réelle verselle d'une matrice réduite à cette forme (avec un nombre minimal de paramètres) si l'on remplace chaque bloc de sa déformation verselle minimale. Le nombre minimal des paramètres d'une déformation verselle réelle est donné par conséquent par la formule

$$d = \sum_{\lambda} [n_1(\lambda) + 3n_2(\lambda) + 5n_3(\lambda) + \ldots],$$

où la sommation est étendue à toutes les ν valeurs propres aussi bien réelles que complexes.

Le travail de Galine comporte des formules explicites pour les déformations verselles et les tables des diagrammes des bifurcations de matrices réelles pour $d - \nu \leqslant 3$. Pour les applications à la mécanique on a dressé des tableaux de déformations verselles de matrices symplectiques et hamiltoniennes (infinitésimalement symplectiques) (il s'agit de déformation préservant la symplecticité) : D. G a l i n e,

Déformation verselle de systèmes hamiltoniens linéaires, Travaux du séminaire de I. Pétrovski, 1975, vyp. 1, pages 103-114 (en russe).

Voici une application des diagrammes des bifurcations obtenus. Supposons qu'en étudiant un phénomène on ait obtenu un diagramme des bifurcations d'une structure autre que celles énumérées. Alors de deux choses l'une: ou bien en simulant le phénomène on a omis un fait important susceptible de modifier qualitativement la structure du diagramme, ou bien il existe des raisons spéciales pour une multiplicité complémentaire du spectre ou pour la non-transversalité à la stratification de Jordan (par exemple, le problème est symétrique ou hamiltonien).

F. Problème de classification des singularités du diagramme du décrément. A titre d'application des déformations verselles des matrices nous allons considérer la résolution du problème suivant. Soit donnée une famille d'équations différentielles autonomes homogènes, linéaires. On sait que l'asymptotique des solutions pour $t \to +\infty$ est définie par la valeur propre de l'opérateur qui définit l'équation possédant la plus grande partie réelle. On demande comment cette partie réelle dépend des paramètres.

En technique la partie réelle indiquée (précédée du signe —) s'appelle décrément, donc notre problème consiste à étudier le comportement du décrément lorsque les paramètres du système varient.

Le comportement du décrément lorsque les paramètres varient est commode à décrire si l'on indique sur le plan (dans l'espace, . . .) les paramètres des lignes (surface, . . .) de niveau du décrément. On appellera diagramme du décrément la famille de lignes de niveau du décrément sur le plan des valeurs des paramètres.

La forme du diagramme du décrément varie beaucoup d'une famille à l'autre; dans certains cas le diagramme du décrément peut présenter des singularités assez complexes. Il s'avère cependant que dans les familles génériques le diagramme du décrément ne présente que certaines singularités simples: les singularités plus complexes se désagrègent par une petite déformation de la famille.

Dans ce paragraphe on décrit toutes les singularités des diagrammes des décréments de familles génériques à deux paramètres.

La classification des singularités du diagramme d'un décrément générique peut rendre lors de l'étude de la dépendance du système par rapport aux paramètres les mêmes services que la classification des points singuliers génériques dans l'étude des trajectoires.

L'apparition d'une singularité non générique sur le diagramme du décrément doit susciter de l'inquiétude: elle peut s'expliquer par une symétrie spéciale de la famille ou bien témoigner d'une simulation inadéquate dans laquelle les petits effets (par exemple les couplages parasites en radiotechnique) négligés dans les équations sont susceptibles de modifier qualitativement les phénomènes.

La classification des singularités des diagrammes des décréments génériques à deux paramètres est utile en particulier pour l'étude des singularités de la frontière du domaine de stabilité dans les familles génériques à trois paramètres d'équations linéaires (surfaces de décréments nuls).

Les résultats obtenus peuvent être appliqués également aux systèmes non linéaires possédant des points stationnaires dépendant différentiablement des paramètres: le décrément du linéarisé du système non linéaire en un tel point possédera comme fonction de paramètres uniquement des singularités simples (dans le cas d'une famille générique).

Lorsqu'on applique les résultats obtenus à des systèmes non linéaires, il faut toutefois exclure la portion de frontière du domaine de stabilité correspondant aux racines nulles, puisque sur cette portion le point stationnaire ne dé-

pend plus différentiablement des paramètres. Donc la description des singularités de la frontière du domaine de stabilité pour les systèmes non linéaires (et la description des diagrammes des décréments au voisinage des points de cette frontière) implique une étude supplémentaire. Nous reviendrons sur cette question dans les paragraphes suivants.

Lorsqu'on étudie les itérations d'applications ainsi que les équations à coefficients périodiques ou les mouvements au voisinage d'une trajectoire périodique, le rôle du décrément est tenu par le plus grand des modules des valeurs propres. Si ce module est différent de l'unité, alors ses singularités (comme fonctions des paramètres dans la famille générique) sont les mêmes que celles du décrément de la famille générique. Pour cette raison dans la suite on n'étudiera que le décrément.

Lorsqu'on étudie les modules des valeurs propres dans des problèmes non linéaires du type indiqué plus haut, on peut se servir des résultats du présent paragraphe hors de la frontière du domaine de stabilité et dans les points de la frontière pour lesquels l'unité n'est pas valeur propre.

G. Diagramme du décrément.

Considérons une famille d'opérateurs linéaires A dans l'espace euclidien $\mathbb{R}^n$, dépendant différentiablement comme d'un paramètre d'un point λ de l'espace des paramètres Λ

$$A(\lambda): \mathbb{R}^n \to \mathbb{R}^n.$$

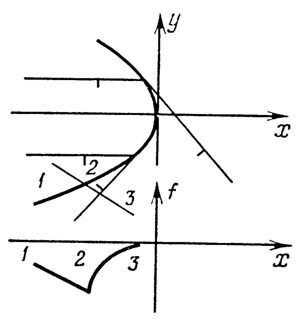


Fig. 113

D é f i n i t i o n. On appelle *incrément* *) d'une famille une fonction f d'un paramètre dont la valeur au point λ est égale à la plus grande partie réelle des valeurs propres de l'opérateur $A(\lambda)$:

$$f(\lambda) = \lim_{t \to +\infty} \frac{1}{t} \ln \| e^{A(\lambda)t} \|.$$

La fonction f est continue mais pas nécessairement différentiable. On se propose d'étudier les singularités des fonctions f pour des familles génériques à deux paramètres. On peut donc admettre que l'espace des paramètres Λ est le plan $\mathbb{R}^2$ ou un domaine du plan.

On appellera diagramme du décrément la famille de lignes de niveau de la fonction f sur le plan. Le trait oblique à la ligne de niveau indiquera la direction de la pente, c'est-à-dire la direction suivant laquelle f diminue. En d'autres termes, le trait est dirigé dans le sens de croissance de la stabilité.

E x e m p l e. Soit l'équation différentielle

$$\ddot{z} = xz + y\dot{z},$$

dépendant de deux paramètres (x, y). La matrice du système correspondant s'écrit

$$A(x, y) = \begin{pmatrix} 0 & 1 \\ x & y \end{pmatrix}.$$

Le diagramme du décrément est représenté sur la figure 113. La parabole $4x + y^2 = 0$ partage le plan (x, y) en deux régions. Dans chacune d'elles l'incrément est une fonction différentiable. A gauche de la parabole les valeurs pro-

*) Les ingénieurs appellent la quantité $|f|$ décrément si $f < 0$ et incrément si $f > 0$.

pres sont complexes et $f = y/2$. A droite les valeurs propres sont réelles et $f = (y + \sqrt{4x + y^2})/2$. Les lignes de niveau de l'incrément sont des rayons parallèles à la parabole.

Tous les points de la parabole sont points singuliers du diagramme du décrément. Il leur correspond des matrices A de bloc de Jordan d'ordre 2. Lorsqu'on coupe la parabole de gauche à droite, le changement linéaire de l'incrément est remplacé par un changement radical. Il est clair que la singularité indiquée ici est inévitable par une petite déformation de la famille. Il existe d'autres singularités inévitables; on se propose de les énumérer toutes.

H. Strates de codimension 1 et 2 dans l'espace des matrices.

Si une seule valeur propre réelle ou un seul couple de valeurs propres *) conjugués complexes de la matrice A (λ_0) possède une partie réelle maximale, alors l'incrément est une fonction différentiable au voisinage de la valeur considérée du paramètre λ_0.

La différentiabilité n'est violée que dans le cas où la valeur propre à partie réelle maximale n'est pas unique. Les matrices dont plusieurs valeurs propres possèdent à la fois une partie réelle maximale forment une sous-variété F semi-algébrique **) fermée dans l'espace $\mathbb{R}^{n^2}$ de toutes les matrices d'ordre n. Cette variété est de codimension 1 et son complémentaire est composé de deux composantes ouvertes:

D_1. *La strate* (α). Une seule valeur propre réelle possède une partie réelle maximale.

D_2. *La strate* $(\alpha \pm i\omega)$. Un couple de valeurs propres conjuguées complexes possède une partie réelle maximale.

La variété F est facilement stratifiable. Voici la liste complète de toutes les strates de dimension maximale (codimension 1):

F_1. *La strate* (α^2). Seules deux valeurs propres confondues possèdent une partie réelle maximale; elles sont réelles et il leur correspond un bloc de Jordan d'ordre 2.

F_2. *La strate* $(\alpha, \alpha \pm i\omega)$. Seules trois valeurs propres possèdent une partie réelle maximale: une valeur réelle et un couple de valeurs conjuguées complexes.

F_3. *La strate* $(\alpha \pm i\omega_1, \alpha \pm i\omega_2)$. Seuls deux couples de valeurs conjuguées complexes distinctes possèdent une partie réelle maximale.

Il est clair que les strates F_1, F_2, F_3 sont des sous-variétés disjointes non fermées, régulières, différentiables de codimension 1 dans l'espace des matrices $\mathbb{R}^{n^2}$. Le reste $F \setminus (F_1 \cup F_2 \cup F_3)$ de la variété F (la variété des matrices possédant plusieurs valeurs propres de partie réelle maximale) est l'adhérence de la sous-variété semi-algébrique fermée de codimension 2 dans l'espace des matrices $\mathbb{R}^{n^2}$. Les strates de dimension maximale de la variété $F \setminus (F_1 \cup F_2 \cup F_3)$ sont de codimension 2 dans $\mathbb{R}^{n^2}$. Voici leur liste:

G_1. *La strate* (α^3). Seules trois valeurs propres possèdent une partie réelle maximale; elles sont réelles et il leur correspond un bloc de Jordan d'ordre 3.

G_2. *La strate* $((\alpha \pm i\omega)^2)$. Seules deux valeurs propres conjuguées complexes confondues possèdent une partie réelle maximale; il leur correspond des blocs de Jordan d'ordre 2.

G_3. *La strate* $(\alpha^2, \alpha \pm i\omega)$. Seules quatre valeurs propres possèdent une partie réelle maximale ; à deux d'entre elles réelles est associé un bloc de Jordan d'ordre 2, les deux complexes forment un couple conjugué complexe.

---

*) Ici et dans la suite on conviendra que les nombres d'un couple conjugué complexe ne sont pas réels.

**) On appelle sous-variété semi-algébrique d'un espace vectoriel la réunion finie des ensembles définis par des systèmes finis d'équations et d'inégalités polynomiales.

G_4. *La strate* $(\alpha, \alpha \pm i\omega_1, \alpha \pm i\omega_2)$. Seules cinq valeurs propres possèdent une partie réelle maximale: une réelle et deux couples conjugués complexes distincts.

G_5. *La strate* $(\alpha \pm i\omega_1, \alpha \pm i\omega_2, \alpha \pm i\omega_3)$. Trois couples conjugués complexes distincts possèdent une partie réelle maximale.

Les strates G_1 à G_5 sont des sous-variétés *) disjointes, non fermées, régulières, de codimension 2 dans l'espace des matrices $\mathbb{R}^{n^2}$. Le reste $F \setminus \bigcup F_i \setminus \bigcup G_i$ est une sous-variété semi-algébrique fermée de codimension 3 dans $\mathbb{R}^{n^2}$.

Le théorème faible de transversalité (§ 29) entraîne le

C o r o l l a i r e. *Les familles génériques de matrices à deux paramètres ne renferment pas de matrices possédant des collections de valeurs propres à partie réelle maximale autres que celles énumérées plus haut* (D_i, F_i, G_i) ; *ces strates ne sont coupées que transversalement par la famille.*

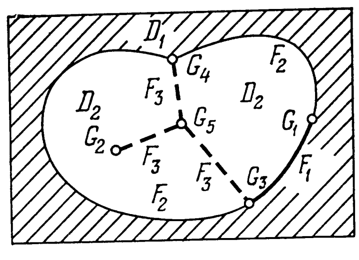


Fig. 114

Donc, dans une famille générique les collections de valeurs propres de codimension 1 (F_i) se rencontrent sur les courbes différentiables dans le plan des paramètres, possédant des points singuliers seulement aux points de ce plan donnant lieu à des collections de codimension 2 (G_i). Cette dernière circonstance ne peut avoir lieu que dans des points isolés du plan des paramètres.

Si l'on ajoute aux segments F_1 et F_2 leurs points singuliers G_i on obtient des courbes qui partagent le plan des paramètres en régions de deux types: D_1 et D_2. Il est aisé de voir que tous les segments F_3 sont contenus dans la région D_2.

Par ailleurs,

les points G_1 (α^3) sont situés sur la jonction de F_1 (α^2) et F_2 $(\alpha, \alpha \pm i\omega)$;

les points G_2 $((\alpha \pm i\omega)^2)$ adhèrent à F_3 $(\alpha \pm i\omega_1, \alpha \pm i\omega_2)$;

les points G_3 $(\alpha^2, \alpha \pm i\omega)$ sont situés sur la jonction de F_1 (α^2), F_2 $(\alpha, \alpha \pm i\omega)$, F_3 $(\alpha \pm i\omega_{1,2})$;

les points G_4 $(\alpha, \alpha \pm i\omega_{1,2})$ sont situés sur la jonction de F_2 $(\alpha, \alpha \pm i\omega)$ et F_3 $(\alpha \pm i\omega_{1,2})$;

les points G_5 $(\alpha \pm i\omega_{1,2,3})$ adhèrent à F_3 $(\alpha \pm i\omega_{1,2})$.

La figure 114 représente un exemple (hypothétique) de configuration que sont susceptibles de former ces lignes sur le plan des paramètres dans une famille générique.

I. Construction du diagramme du décrément au voisinage des points des strates de codimension 0 et 1.

L'incrément f est une fonction différentiable des paramètres sur le complémentaire à l'ensemble des singularités de F. Cependant le diagramme du décrément peut posséder des singularités en certains points de ce complémentaire: les points critiques de la fonction f.

En dehors de F l'incrément de la famille générique ne possède que des points critiques simples, c'est-à-dire des points des trois types suivants qui se transforment en six types si l'on distingue les cas des racines réelles (D_1) et complexes (D_2):

---

*) Toutes les variétés D_i, F_i, G_i sont connexes pour n assez grands. Exceptions: D_2 et F_1 pour $n=2$, F_3, G_2, G_3 pour $n = $ ′, et G_5 pour $n = 6$, possèdent chacune deux composantes.

D_i^0. *Minimum*. Au voisinage du point considéré du plan des paramètres on peut choisir des coordonnées différentiables x, y, tels que l'incrément soit de la forme $f = \text{const} + x^2 + y^2$.

D_i^1. *Col*. $f = \text{const} + x^2 - y^2$ dans un système de coordonnées convenables.

D_i^2. *Maximum*. $f = \text{const} - x^2 - y^2$.

Etudions maintenant la fonction f au voisinage des points non singuliers de l'ensemble F, c'est-à-dire au voisinage des points intérieurs des courbes F_i du plan des paramètres. On distinguera deux cas selon que le point de la courbe F_i est critique ou non pour l'incrément traité comme une fonction différentiable sur cette courbe.

Le théorème de transversalité entraîne que, dans les familles génériques, les points critiques des restrictions de l'incrément aux courbes F_i ne peuvent être que des maximums ou des minimums non dégénérés.

En combinant cette information et les formules explicites des familles verselles de matrices du point D, on est conduit sans peine aux formes normales suivantes de l'incrément au voisinage des points des strates de codimension un.

Théorème. *Au voisinage d'un point non critique de la restriction de l'incrément d'une famille générique à une courbe F_i, on peut choisir des coordonnées différentiables (x, y) sur le plan des paramètres dans lesquelles l'incrément f prend l'une des trois formes suivantes* (fig. 115) :

Cas F_1^0 (bloc de Jordan) :

$$f = \text{const} + y + \begin{cases} \sqrt{x}, & si\ x \geqslant 0, \\ 0 & si\ x \leqslant 0. \end{cases}$$

Cas F_2^0 et F_3^0 (dépassement simple) :

$$f = \text{const} + x + |y|.$$

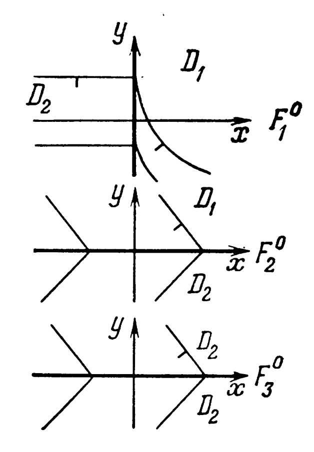


Fig. 115

Les courbes F_1^0 et F_2^0 partagent les domaines des racines réelles et complexes D_1 et D_2. Les lignes de niveau de l'incrément présentent un contact avec la courbe F_1 du côté des racines réelles et lui sont transversales du côté des racines complexes. Les lignes de niveau du décrément abordent les points F_2^0 et F_3^0 des courbes F_2 et F_3 transversalement des deux côtés. L'angle convexe des lignes brisées de niveau aux points de chaque ligne F_i contient dans tous les cas la direction de décroissance de f le long de cette ligne.

Théorème. *Au voisinage d'un point critique de la restriction de l'incrément d'une famille générique on peut choisir des coordonnées (x, y) dans lesquelles l'incrément prend l'une des 12 formes énumérées plus bas* (fig. 116).

Cas F_2^k et F_3^k, $k = 1, \ldots, 4$ (extrémum lié pendant le dépassement) :

$$f = \text{const} + \varepsilon x^2 + \varphi(y) + |y|, \quad \varepsilon = (-1)^k,$$

où $\varphi(y) = ay + \ldots$ est une fonction différentiable, $a > 0$, $a \neq 1$.

On obtient quatre valeurs de k en combinant les deux signes de ε et les deux variantes de a:

k	1,2	3,4
	(0,1)	$(1, +\infty)$

Les k impairs correspondent à un maximum lié, les pairs, à un minimum. Pour obtenir une représentation explicite de la forme du diagramme du décrément il suffit de considérer le cas $\varphi(y) = ay$: dans ce cas les lignes de niveau de f sont composées de portions de deux paraboles déplacées le long de l'axe y.

Cas F_1^k, $k = 1, \ldots, 4$ (extrémum lié de bloc de Jordan α^2):

$$f = \text{const} + \varepsilon x^2 + \varphi(y) + \begin{cases} \sqrt{y} & \text{si } y \geqslant 0, \\ 0 & \text{si } y \leqslant 0, \end{cases}$$

où $\varepsilon = \pm 1$, $\varphi(y) = ay + \ldots$ est une fonction différentiable, $a \neq 0$.

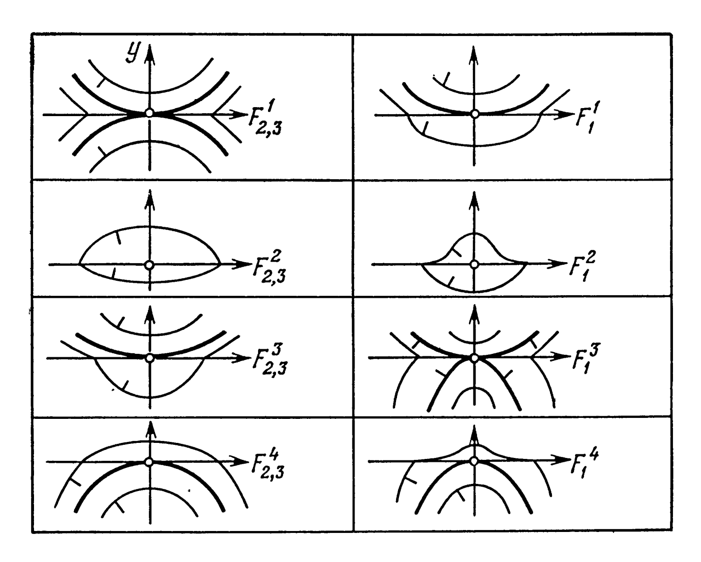


Fig. 116

On obtient quatre valeurs de k en combinant les signes de ε et a:

k	1	2	3	4
signe de ε, signe de a	$--$	$-+$	$+-$	$++$

Les k impairs correspondent à une maximum lié, les pairs, à un minimum. Pour obtenir la forme explicite du décrément il suffit de considérer le cas $\varphi(y) = \pm y$.

Notre théorème affirme que l'incrément de la famille générique à deux paramètres ne possède aucune singularité en les points intérieurs des courbes F autres que celles des 15 types F_i^k ($15 = 3 + 12$) énumérés: si une famille comporte d'autres singularités, on peut les éliminer par une déformation aussi petite que l'on veut de cette famille. Les singularités des types F_i^k sont visiblement inévitables

J. Construction du diagramme du décrément au voisinage de strates de codimension 2.

Lorsqu'on étudie les singularités de strates de codimension 2 dans des familles génériques à deux paramètres, on peut considérer seulement les cas les plus

« non dégénérés », puisque toute dégénérescence accroît la codimension et la singularité devient artificielle.

En combinant le théorème de transversalité et les formules explicites pour les familles verselles de matrices du point B, on est conduit aux formes normales suivantes de l'incrément au voisinage des points de strates de codimension 2.

Théorème. *Au voisinage d'un point de chaque strate de codimension* 2 (G_i *dans les notations du point* H) *on peut choisir sur le plan des paramètres d'une*

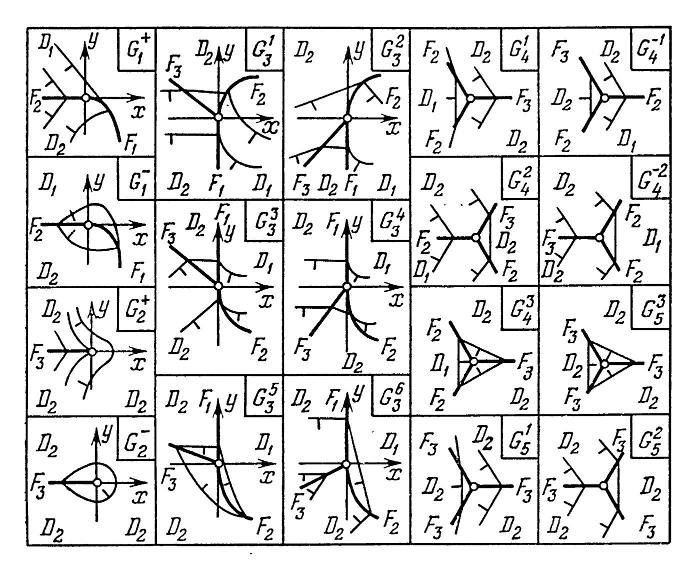


Fig. 117

famille générique des coordonnées différentiables (x, y) *dans lesquelles l'incrément* f *prend l'une des* 18 *formes énumérées plus bas* (fig. 117).

Cas $G_1^{\pm}$ (bloc de Jordan d'ordre 3):

$$f = \varphi(x, y) + \lambda(x, y),$$

où λ est la plus grande des parties réelles des racines de l'équation cubique $\lambda^3 = x\lambda + y$, et φ est une fonction différentiable, telle que $(\partial\varphi/\partial x)(0, 0) = a \neq 0$.

La forme du diagramme du décrément est déterminée par le signe de a.

Le signe « + » ou « — » dans $G_1^{\pm}$ correspond à $a > 0$ et à $a < 0$ respectivement. Pour avoir une idée exacte de la forme du diagramme du décrément il suffit de considérer les cas $\varphi = \pm x$. Deux courbes singulières présentent un contact au point $x = y = 0$: le rayon F_2 ($y = 0$, $x < 0$) et la moitié de la parabole semi-cubique F_1 ($4x^3 = 27y^2$, $y < 0$). Ces deux courbes séparent le domaine des racines conjugués complexes D_2 (convexe) du domaine D_1 des racines réelles. Lorsqu'on se déplace le long de la frontière des domaines D_1 et D_2, l'incrément f varie de façon monotone pour $a > 0$ et présente un minimum au point G_1^- pour $a < 0$. Les lignes de niveau de f sont tangentes à la parabole semi-cubique F_1 du côté de D_1.

Cas $G_2^{\pm}$ (couple complexe de 2-blocs de Jordan):

$$f = \varphi(x, y) + |\mathrm{Re}\sqrt{x + iy}|,$$

où Re est la partie réelle, φ une fonction différentiable telle que $(\partial\varphi/\partial x)(0, 0) = a \neq 0$. La forme du diagramme du décrément est définie par le signe du nombre a.

Les signes « + » et « — » dans $G_2^{\pm}$ correspondent à $a > 0$ et à $a < 0$ respectivement. Pour avoir une idée claire de la forme du diagramme du décrément il suffit de traiter le cas $\varphi = \pm x$. Le rayon F_3 ($y = 0$, $x < 0$) s'approche (et aboutit) au point $x = y = 0$. Pour $a < 0$ la fonction f présente un minimum au point G_2^- ($x = y = 0$). Pour $a > 0$ le point G_2^+ ($x = y = 0$) est topologiquement non singulier pour la fonction f. En passant par ce point, la ligne de niveau de la fonction f présente une singularité de type semi-cubique.

Cas G_3^k ($k = 1, \ldots, 6$; collision d'un couple complexe à bloc de Jordan):

$$f = \text{const} + y + \max \begin{cases} \sqrt{x}, \ \varphi(x, y) & \text{si } x \geqslant 0, \\ 0, \ \varphi(x, y) & \text{si } x \leqslant 0, \end{cases}$$

où $\varphi(x, y) = ax + by + \ldots$ est une fonction différentiable, $a \neq 0$, $b \neq 0$ et $b \neq -1$.

On obtient six valeurs de k en combinant les signes de a et les trois intervalles de variation de b:

k	1	2	3	4	5	6
signe de a	+	—	—	+	—	+
intervalle de b	$(0, +\infty)$	$(0, +\infty)$	$(-1, 0)$	$(-1, 0)$	$(-\infty, -1)$	$(-\infty, -1)$

Pour avoir une idée nette de la forme du diagramme du décrément il suffit de prendre la fonction φ linéaire. Trois rayons différentiables F_1, F_2, F_3 abordent le point ($x = y = 0$), F_1 et F_2 viennent à la rencontre l'un de l'autre (avec un contact d'ordre 1), F_3 se rapproche transversalement du côté du domaine D_2 des racines complexes. Dans le cas G_3^5 (c'est-à-dire $a < 0$, $b < -1$) l'incrément possède au point $x = y = 0$ un minimum; dans les autres cas, le point G_3^k ($k \neq 5$) est un point topologiquement non singulier de la fonction f.

Cas G_5^k ($k = 1, 2, 3$) (double dépassement):

$$\varphi = \text{const} + x + \max(|y|, \varphi(x, y)),$$

où $\varphi(x, y) = ax + by + \ldots$ est une fonction différentiable, $a < 0$, $b > 0$, $a + 1 \neq \pm b$.

Les trois valeurs de k pour G_5^k correspondent aux intervalles de variation de a:

k	1	2	3
condition sur a	$b - 1 < a$	$-b - 1 < a < b - 1$	$a < -b - 1$

Pour avoir une idée claire de la forme du diagramme du décrément il suffit de prendre la fonction φ linéaire.

Dans chacun des trois cas ($k = 1, 2, 3$) trois branches différentiables de la courbe F_3 convergent transversalement au point G_5^k. Dans le dernier cas, ce point est un minimum de l'incrément, dans les deux premiers un point topologiquement non singulier. A l'abord du point G_5^k l'incrément décroît sur k des trois rayons et croît sur les autres.

Cas G_4^k ($k = \pm 1, \pm 2, 3$) (double dépassement avec la participation d'une racine réelle). L'incrément est donné par la même formule que dans les cas G_5^k, mais il faut distinguer plus de variantes selon que la racine réelle correspond à tel ou tel vecteur .

Aux k négatifs correspondent les cas dans lesquels le décrément croît sur la ligne F_3 (sur laquelle a lieu la collision de couples complexes) à l'approche du point G_4^k. Les deux autres rayons sont des branches de la courbe F_2.

K. Discussion.

L'étude des formes normales énumérées ci-dessus nous permet de tirer des conclusions générales sur la structure du diagramme du décrément tant localement que globalement. Une première conséquence de nos théorèmes est le

C o r o l l a i r e. *L'incrément $f: \Lambda \to \mathbb{R}$ d'une famille générique à deux paramètres est topologiquement équivalent à une fonction différentiable ne possédant que des points critiques simples.*

Plus exactement, les points de minimum sont des points de types D_i^0, F_i^2, $G_{1,2}^-$, G_3^5, $G_{4,5}^3$.

Les points D_i^1 et F_i^3 sont topologiquement équivalents à un col. L'incrément est une fonction différentiable au voisinage des points de maximum (D_i^2). Les points de tous les autres types sont topologiquement non singuliers.

Ce corollaire nous permet de toute évidence d'établir des inégalités pour le nombre de points singuliers des divers types. En particulier *si une ligne de niveau fermée quelconque de l'incrément est la frontière d'un domaine simplement connexe, alors le nombre total de points de types $D_i^{0,2}$, F_i^2, $G_{1,2}^-$, G_3^5, $G_{4,5}^3$, à l'intérieur de ce domaine, est supérieur de* 1 *au nombre de points D_i^1, F_i^3*. On ne sait pas si cette assertion se généralise à des familles l-paramétriques, $l > 2$ *).

Du fait que les portions F_1 et F_2 forment des courbes fermées et de la description des singularités aux extrémités de F_3 il résulte le

C o r o l l a i r e. *Si l'espace des paramètres Λ est une variété fermée à deux dimensions, les nombres de points de types G_1 et G_3 ont même parité, et le nombre total des points de types G_2, G_3, G_4, G_5 est pair.*

Si pour espace des paramètres Λ on prend un domaine compact à bord coupant transversalement F_i et ne passant pas par les points G_i, alors le résultat est modifié comme suit : le nombre total de points de types G_1 et G_3 est de même parité que le nombre total de points d'intersection du bord avec F_1 et F_2, et le nombre total de points de types G_2, G_3, G_4, G_5, de même parité que le nombre de points d'intersection du bord avec F_3.

Cette étude de l'incrément nous permet, en particulier, d'envisager les singularités de la frontière de stabilité (c'est-à-dire de la ligne des incréments nuls) sur le plan des paramètres de familles génériques à deux paramètres. Nos théorèmes entraînent le

C o r o l l a i r e. *La frontière de stabilité d'une famille générique à deux paramètres de matrices est composée d'arcs différentiables se coupant transversalement en leurs extrémités.*

On remarquera que les points anguleux de la frontière de stabilité peuvent être d'après la classification des points I et J de types F_1^0 (« 2-bloc de Jordan ») ou F_2^0, F_3^0 (« dépassement simple »). Chaque arc de frontière de stabilité est prolongeable, de ce fait, au-delà de ses extrémités sans perte de différentiabilité. Ceci étant le nombre total de points anguleux de types F_1^0 et F_2^0 sur chaque composante fermée de la frontière de stabilité est toujours pair.

On remarquera par ailleurs que l'analyse des singularités de l'incrément de familles à deux paramètres est suffisante pour étudier la frontière de stabilité dans les familles à trois paramètres.

En effet, le théorème de transversalité nous dit que l'on peut éliminer les

*) On remarquera que pour $l = 2$ les singularités de l'incrément de la famille générique sont les mêmes que celles de la plus grande partie réelle d'une racine d'un polynôme algébrique dont les coefficients sont des fonctions génériques de l paramètres. Pour $l \geqslant 3$ il en va autrement : l'incrément peut présenter des singularités plus complexes.

points singuliers des strates de codimension 3 et les points critiques des restrictions de l'incrément à des strates de codimension 0,1 et 2 de la frontière de stabilité par une petite déformation de la famille. Donc la frontière de stabilité de la famille générique est constituée de surfaces différentiables, et ses singularités sont situées sur les courbes suivant lesquelles la frontière de stabilité coupe les surfaces de types F_i et dans les points d'intersection de la frontière de stabilité avec les strates G_i (ces dernières sont des courbes dans les familles génériques à trois paramètres).

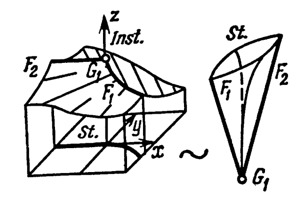
Si l'on se déplace le long d'une telle courbe G_i, on peut considérer la famille à trois paramètres comme une famille à un paramètre de familles à deux paramètres (deux paramètres sont les coordonnées dans le plan transversal à G_i et un paramètre est la coordonnée t le long de G_i). En étudiant les formes normales des points I et J, on doit considérer maintenant toutes les constantes et les fonctions arbitraires φ dépendant différentiablement du paramètre t. Bien plus dans le cas générique nous pouvons prendre pour paramètre z les fonctions φ (x, y, t). On est conduit de ce fait à la conclusion suivante.

C o r o l l a i r e. Les singularités de la frontière de stabilité d'une famille générique à trois paramètres de matrices sont les mêmes que celles des graphes des incréments des familles génériques à deux paramètres *). Voici la liste complète de ces singularités à un difféomorphisme près **) (fig. 118) :

Le dièdre (F_i) : $|y| + z = 0$.

Le trièdre $(G_{3,4,5})$: $z + \max(x, |y|) = 0$.

Impasse sur l'arête (G_2) : $z + |\operatorname{Re}\sqrt{x+iy}| = 0$. (*Cette surface est difféomorphe dans* $\mathbb{R}^3$ *à une surface définie par l'équation* $XY^2 = Z^2$, *où* $X \geqslant 0$, $Y \geqslant 0$.)



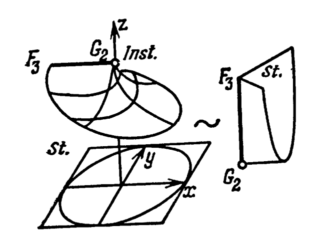


Fig. 118

Le coude de l'arête (G_1) : $z + \lambda(x, y) = 0$, *où* λ *est la plus grande des parties réelles des racines de l'équation* $\lambda^3 = x\lambda + y$. (*Cette surface est difféomorphe dans* $\mathbb{R}^3$ *à la surface définie par l'équation* $X^2Y^2 = Z^2$, *où* $X \geqslant 0$, $Y \geqslant 0$.)

Les angles de la frontière de stabilité ont toujours leurs sommets à l'extérieur (ils forment des enclaves dans le domaine d'instabilité). Ceci est probablement une illustration du principe général qui veut que le bien soit fragile.

Certaines propriétés globales de la frontière de stabilité découlent de ce qui vient d'être dit. Si, par exemple, la frontière de stabilité est fermée, le nombre total de sommets de types $(G_i, i > 1)$ est pair de même que le nombre de sommets de types G_1 et G_3.

On trouvera la démonstration de ces théorèmes dans l'article de V. A r n o l d, Cours sur les bifurcations et les familles verselles. UMN, **27**, 5 (1972), 119-184, en russe.

---

*) Pour des raisons analogues les singularités de la frontière de stabilité des familles à $n + 1$ paramètres sont les mêmes que celles des graphes d'incrément à n paramètres.

**) Il s'agit d'une application prolongeable à un difféomorphisme du voisinage de la surface.

§ 31. Bifurcation des points singuliers d'un champ de vecteurs

Dans ce paragraphe on étudie des familles à un paramètre d'équations différentielles et les bifurcations des points singuliers de familles génériques.

A. Courbe de points singuliers.

Soit un champ de vecteurs dépendant différentiablement d'un paramètre. Supposons que ce champ possède un point singulier pour une certaine valeur du paramètre. Que deviendra ce point singulier si le paramètre varie?

Théorème. *Un point singulier d'un champ de vecteurs, dépendant différentiablement d'un paramètre, dépend lui-même différentiablement de ce paramètre tant que toutes les valeurs propres de la partie linéaire de ce champ sont distinctes de zéro en ce point singulier.*

◀ La famille de champs est donnée par n fonctions de $n + 1$ variables (n coordonnées et un paramètre ε) au voisinage du point étudié et de la valeur du paramètre étudiée dans un espace des phases à n dimensions. Les points singuliers sont définis par un système de n équations en $n + 1$ variables, $v(x, \varepsilon) = 0$. Le théorème des fonctions implicites nous apprend que ces équations définissent localement une courbe différentiable $x = \gamma(\varepsilon)$ si le déterminant de $\partial v/\partial x$ est différent de 0 au point initial considéré. Or, ce déterminant est égal au produit des valeurs propres du linéarisé du champ au point singulier. Il est nul par hypothèse. ▶

Remarque. Les points singuliers en lesquels les valeurs propres du champ linéarisé sont différentes de zéro sont dits *non dégénérés.* Donc, si le champ dépend différentiablement d'un paramètre, ses points singuliers dépendent différentiablement de ce paramètre tant qu'ils sont non dégénérés. Cette démonstration est valable pour un espace de paramètres de dimension quelconque.

Tous les points singuliers d'un champ générique sont non dégénérés. Cependant si l'on étudie une famille de champs de vecteurs, pour certaines valeurs du paramètre il peut se présenter des singularités inévitables par une petite déformation de la famille.

Etudions les singularités des familles génériques à un paramètre de champs de vecteurs dans un espace à n dimensions.

Considérons l'espace à $n + 1$ dimensions, produit direct de l'espace des phases par l'axe des valeurs du paramètre ε. On désignera par x un point de l'espace des phases. Notre famille définit une famille d'équations différentielles

$$\dot{x} = v(x, \varepsilon).$$

Considérons dans l'espace à $n+1$ dimensions l'ensemble des points singuliers des équations de la famille pour toutes les valeurs du paramètre (fig. 119):

$$\Gamma = \{x,\ \varepsilon : v(x,\ \varepsilon) = 0\}.$$

Théorème. *L'ensemble des points singuliers d'une famille générique est une courbe différentiable.*

Ici et dans la suite par « familles génériques » on comprendra des « familles d'un ensemble partout dense dans l'espace de toutes

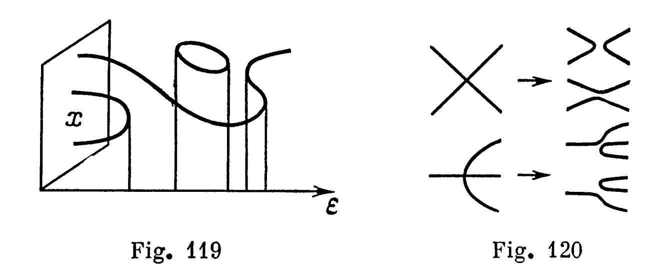


Fig. 119 Fig. 120

les familles »; l'ensemble partout dense en question est ouvert si le domaine de définition de la famille est compact ou si les familles sont considérées pour une topologie fine (cf. § 29); en tous les cas cet ensemble partout dense est l'intersection d'un nombre dénombrable d'ensembles ouverts.

◀ Le théorème est une conséquence du théorème de transversalité (§ 29) ou du lemme de Sard (§ 10).

En effet, suivant le théorème des fonctions implicites, Γ est une courbe localement différentiable si 0 n'est pas valeur critique de l'application locale $\mathbb{R}^{n+1} \to \mathbb{R}^n$, $(x,\ \varepsilon) \mapsto v(x,\ \varepsilon)$. Or, la valeur 0 n'est pas critique pour une application générique.

Remarque. Dans ce théorème la dimension de l'espace des paramètres n'est pas non plus essentielle.

Le théorème démontré exclut immédiatement certaines bifurcations des points singuliers.

Considérons, par exemple, les bifurcations représentées sur la figure 120 à gauche. Le théorème nous dit que ces bifurcations ne sont pas préservées par une petite déformation de la famille.

En effet, il est aisé de voir que ces bifurcations se désagrègent par une petite déformation avec l'un des procédés indiqués sur la figure 120 à droite. Si dans un problème on est confronté à des bifurcations de la forme indiquée sur la figure 120 à gauche, cela veut dire que la famille considérée n'est pas générique. Ceci traduit une symétrie spéciale de la situation ou encore une idéalisation inadé-

quate négligeant de petits effets susceptibles toutefois de modifier qualitativement le comportement des points singuliers en fonction du paramètre. Pour établir lequel de ces cas se présente dans le système réel dont l'idéalisation a donné lieu à cette bifurcation non générique, il faut calculer certains des termes de l'équation différentielle négligés. Les formules qui suivront nous indiquent quels termes il importe de calculer.

B. Valeurs de bifurcation du paramètre.

Supposons que l'ensemble des points singuliers de la famille est une courbe différentiable (rang $(\partial v/\partial\,(x,\ \varepsilon)) = n$). Considérons la projection de cette courbe différentiable sur l'axe des valeurs du paramètre. Les points en lesquels cette courbe se projette mal sur l'axe ε sont précisément les points singuliers dégénérés. En effet, d'après le théorème des fonctions implicites la courbe des points singuliers est le graphe d'une fonction différentiable du paramètre au voisinage du point singulier non dégénéré.

Définition. On appelle *valeur de bifurcation du paramètre* la valeur de ce paramètre correspondant à un point singulier dégénéré, et *point de bifurcation*, le point singulier dégénéré dans le produit direct de l'espace des phases par l'axe de valeurs du paramètre.

Traitons la valeur du paramètre ε comme une fonction sur la courbe des points singuliers. *Les valeurs de bifurcation du paramètre sont les valeurs critiques de cette fonction, les points de bifurcation, les points critiques de la fonction* (les points où la différentielle de la fonction est nulle).

Un point critique d'une fonction est *non dégénéré* si la deuxième différentielle de la fonction est non dégénérée en ce point (comme nous avons affaire à des fonctions d'une variable, la non-dégénérescence de la deuxième différentielle signifie simplement sa non-nullité). Le point de bifurcation correspondant est appelé *point de bifurcation non dégénéré.*

Définition. La valeur de bifurcation d'un paramètre est *régulière* s'il lui correspond exactement un point de bifurcation, et ce point est non dégénéré.

Théorème. *Toutes les valeurs de bifurcation du paramètre sont régulières pour les familles génériques à un paramètre. Si l'espace des phases est compact, les valeurs de bifurcations sont isolées.*

◀ Cette proposition est une conséquence immédiate du théorème de transversalité. La démonstration est laissée au soin du lecteur. ▶

Remarque. Le théorème veut dire qu'au fur et à mesure que le paramètre varie, les points singuliers de la famille générique ne peuvent que se détruire deux à deux ou naître par couple lorsque le paramètre prend une valeur de bifurcation (fig. 119). Les bifurcations de ce type sont stables (sont préservées par une petite défor-

mation de la famille). Les bifurcations plus compliquées se désagrègent en plusieurs bifurcations du type décrit (fig. 121) par une petite déformation de la famille.

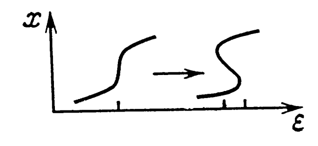


Fig. 121

C. Exemple. Champ de vecteurs sur une droite.

Soit une famille à un paramètre de champs de vecteurs sur une droite définissant l'équation différentielle

$$\dot{x} = \pm x^2 + \varepsilon, \quad x \in \mathbb{R}, \quad \varepsilon \in \mathbb{R}.$$

Pour $\varepsilon = 0$ ce champ de vecteurs présente un point singulier dégénéré simple ($x = 0$). Le passage du paramètre par une valeur de bifurcation régulière $\varepsilon = 0$ donne lieu en fonction du signe de x^2 soit à une élimination mutuelle de deux points singuliers, un point stable et un point instable, soit à la naissance d'un couple de points singuliers qui divergent aussitôt (avec une asymptotique $\sqrt{|\varepsilon|}$).

On vérifie sans peine que la bifurcation de cet exemple est l'unique bifurcation inévitable dans les familles génériques à un paramètre de champs vecteurs sur une droite.

D é f i n i t i o n. Soient données deux familles de champs de vecteurs dépendant d'un paramètre. Ces familles sont dites *topologiquement équivalentes* si existent un homéomorphisme des espaces des paramètres et une famille d'homéomorphismes de l'espace des phases dépendant continûment du paramètre et envoyant les trajectoires orientées d'une famille pour chaque valeur du paramètre dans la famille des trajectoires orientées de la deuxième famille pour la valeur correspondante du paramètre.

A noter que ces homéomorphismes définissent un homéomorphisme des produits directs des espaces des phases et des espaces des valeurs des paramètres $(x, \varepsilon) \mapsto (h(x, \varepsilon), \varphi(\varepsilon))$, envoyant les trajectoires du système $\dot{x} = v(x, \varepsilon)$, $\dot{\varepsilon} = 0$ dans des trajectoires d'un second système de la même forme.

On définit de façon analogue l'équivalence des germes de familles en un point. Si le couple (x_0, ε_0) est constitué d'un point de l'espace des phases et d'un point de l'espace des paramètres, les homéomorphismes réalisant l'équivalence doivent définir un homéomorphisme $(x, \varepsilon) \mapsto (h(x, \varepsilon), \varphi(\varepsilon))$ d'un voisinage du point (x_0, ε_0) dans le produit direct.

T h é o r è m e. *Une famille à un paramètre de champs de vecteurs*

sur une droite est équivalente, au voisinage d'un point de bifurcation non dégénéré, au germe de la famille définissant l'équation $\dot{x} = x^2 + \varepsilon$ *au point* $x = 0$, $\varepsilon = 0$.

◀ La fonction $v(x, \varepsilon)$ définissant le champ change de signe sur la courbe Γ. Plaçons l'origine des coordonnées (x, ε) au point de bifurcation. Ce point étant non dégénéré, l'équation de la courbe Γ s'écrit $\varepsilon = Cx^2 + O(|x|^3)$, $C \neq 0$. D'où notre assertion. ▶

Ce théorème et celui du point B nous donnent une description topologique complète des bifurcations des points singuliers des champs de vecteurs sur une droite dans des familles génériques à un paramètre.

D. Bifurcations de solutions périodiques

On étudie de façon analogue les bifurcations des points invariants d'applications différentiables et les bifurcations des solutions périodiques des équations différentielles (bifurcations des trajectoires fermées ou des courbes intégrales). La condition de non-dégénérescence d'un point fixe d'une application consiste en ce que toutes les valeurs propres de la linéarisée soient distinctes de l'unité. Si les solutions sont périodiques, les valeurs propres de la linéarisée de l'application de Poincaré (c'est-à-dire les valeurs propres de l'opérateur de monodromie défini par une équation aux variations normales le long de la solution considérée) doivent être distinctes de l'unité.

En particulier, si l'équation $\dot{x} = v(x, \varepsilon)$ pour $\varepsilon = 0$ possède une solution périodique $x = \varphi(t)$ de période T et si est réalisée la condition de non-dégénérescence indiquée ci-dessus, alors pour ε petit il existe une solution périodique unique $x = \Phi(t, \varepsilon)$ de période $T(\varepsilon)$, confondue avec φ pour $\varepsilon = 0$ (il est entendu que c'est la trajectoire qui est unique: l'origine des temps peut être variée).

R e m a r q u e. La recherche de la solution périodique Φ sous la forme d'une série sur ε s'appelle *méthode du petit paramètre* de Poincaré; la solution φ est appelée *solution génératrice*. Une méthode analogue est utilisée dans le cas non autonome, où v possède une période $T(\varepsilon)$ en t et où l'on cherche les solutions $T(\varepsilon)$-périodiques.

E x e r c i c e. Trouver à ε^2 près la solution 2π-périodique de l'équation $\ddot{x} = \sin x + \varepsilon \cos t$, qui se confond avec $x \equiv 0$ pour $\varepsilon = 0$.

§ 32. Déformations verselles des portraits de phases

Dans ce paragraphe on définit les déformations topologiquement verselles des portraits de phases *) et on les explicite pour les points singuliers dégénérés simples.

*) On appelle portrait de phases (d'un champ de vecteurs ou d'un système d'équations différentielles) la décomposition de l'espace des phases en trajectoires du champ correspondant.

A. Théorie des bifurcations locales et théorie qualitative locale. On a vu plus haut que les points singuliers dégénérés sont inévitables lorsqu'on s'intéresse non pas à un champ de vecteurs mais à une famille de champs dépendant d'un paramètre. En outre les familles génériques ne donnent lieu qu'à des dégénérescences simples.

Lorsqu'on étudie la structure d'un champ de vecteurs au voisinage d'un point singulier dégénéré, on peut appliquer les méthodes ordinaires de la théorie qualitative des équations différentielles (cf. chap. 3). Dans le cas de singularité simple ces méthodes permettent de procéder à une étude topologique assez complète du portrait de phases. Nous sommes donc en mesure d'étudier le portrait de phases local aussi bien pour les valeurs génériques du paramètre que pour les valeurs singulières. Telle est l'approche habituelle des problèmes relatifs aux familles d'équations différentielles.

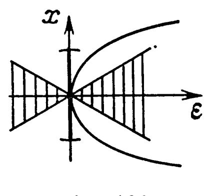


Fig. 122

L'étude des bifurcations simples montrent que cette approche néglige l'essence même des phénomènes qui ont lieu au voisinage d'un point critique du paramètre. En effet, le rayon du voisinage d'un point singulier non dégénéré en lequel le portrait de phases est donné par la théorie locale tend vers 0 à l'approche d'une valeur singulière du paramètre (fig. 122) et croît de nouveau par un saut pour cette valeur singulière du paramètre. De ce fait la modification du portrait de phases (disons le rapprochement d'un point singulier voisin) n'entre pas dans le domaine d'application de la théorie locale.

Donc, la théorie locale omet le phénomène le plus important qui se déroule pour une valeur singulière du paramètre: le phénomène de la bifurcation.

Ainsi donc, *l'étude des points singuliers dégénérés ne présente un intérêt réel que dans le cas où elle va de pair avec celles des familles dans lesquelles la singularité considérée est inévitable, et de plus au voisinage du point singulier dégénéré dans le produit direct de l'espace des phases par l'espace du paramètre*. En d'autres termes, le voisinage du point singulier de l'espace des phases dans lequel il faut étudier le portrait de phases ne doit pas dépendre du paramètre (ne doit pas se réduire à 0 lorsque le paramètre se rapproche d'une valeur singulière).

Des raisonnements absolument identiques montrent combien il est dangereux de ne pas définir exactement le nombre de paramètres essentiels à l'étude de la bifurcation considérée. Par exemple, lorsqu'on étudie un phénomène à deux paramètres essentiels du point de vue « à un paramètre », un phénomène ordinaire sera le phénomène suivant (fig. 123) : pour chaque valeur du paramètre négligé on réussit à étudier la bifurcation dans une famille à un paramètre d'équations dépendant du second paramètre. Or l'intervalle des valeurs du

second paramètre au voisinage de la valeur singulière dans laquelle on réussit à effectuer cette étude tendra vers 0 à l'approche du paramètre négligé de la valeur singulière. La considération du problème comme un problème à deux paramètres (c'est-à-dire dans un voisinage de la valeur singulière du premier paramètre, ne dépendant pas de la valeur du second) permet d'étudier par les méthodes locales la bifurcation qui du point de vue « à un paramètre » sont globales.

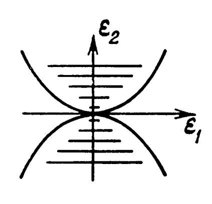


Fig. 123

La perte de stabilité d'une trajectoire fermée est un exemple de problèmes à deux paramètres donnant l'impression d'être de prime abord à un paramètre. Le paramètre naturel est ici le module de la valeur propre de l'opérateur de monodromie; le second paramètre qui est habituellement négligé est l'argument de la valeur propre qui traverse le cercle unitaire. Nous reviendrons sur cet exemple au § 34.

B. Déformations topologiquement verselles.

Soit la famille d'équations différentielles $\dot{x} = v(x, \varepsilon)$.

On appelle *famille locale* $(v;\ x_0, \varepsilon_0)$ le germe de l'application v au point (x_0, ε_0) du produit direct de l'espace des phases par l'espace des paramètres. Donc, chaque représentant de ce germe est défini au voisinage du point (x_0, ε_0) du produit direct (et non pas dans le voisinage du point x_0 de l'espace des phases).

On appelle *équivalence* des familles locales (1) $(v;\ x_0, \varepsilon_0)$ et (2) $(w;\ y_0, \varepsilon_0)$ le germe (au point (x_0, ε_0)) d'une application continue h, $y = h(x, \varepsilon)$, pour les représentants duquel $h(\cdot, \varepsilon)$ est pour chaque ε un homéomorphisme envoyant les trajectoires du système (1) (dans le domaine de définition de h) dans les trajectoires du système (2) avec conservation du sens du mouvement, et de plus $h(x_0, \varepsilon_0) = y_0$. A noter que l'image du point x_0 par l'application $h(\cdot, \varepsilon)$ n'est pas nécessairement le point y_0.

La famille locale (3) $(u;\ x_0, \mu_0)$ *est induite* par la famille (1), à l'aide du germe de l'application continue φ, $\varepsilon = \varphi(\mu)$, au point μ_0, où $\varphi(\mu_0) = \varepsilon_0$ si $u(x, \mu) = v(x, \varphi(\mu))$.

La famille locale (v, x_0, ε_0) s'appelle *déformation verselle du germe du champ* $v_0 = v(\cdot, \varepsilon_0)$ *au point* x_0 si toute autre famille locale renfermant le même germe est équivalente à une famille induite par (v, x_0, ε_0).

Dans la suite, on parlera parfois des déformations, des équivalences, des déformations induites et verselles d'équations différentielles comme synonymes des notions correspondantes pour les champs de vecteurs définissant ces équations.

A signaler que l'existence d'une déformation topologiquement verselle d'un germe donné d'un champ de vecteurs est loin d'être

évidente ; il n'est pas difficile d'exhiber des exemples de champs ne permettant pas une telle déformation avec un nombre fini de paramètres (par exemple, le champ nul). Cependant dans les cas où la déformation verselle existe, a été trouvée et étudiée, l'information qu'elle fournit est très importante. L'indication et l'étude de la déformation verselle constituent une méthode de représentation condensée des résultats d'un examen complet des bifurcations des portraits de phases.

E x e m p l e. La déformation $\dot{x} = \pm x^2 + \varepsilon$ de l'équation différentielle $\dot{x} = \pm x^2$ est verselle.

◀ Voir le § précédent. ▶

C. Théorème de réduction de Chochitaïchvili.

La bifurcation de l'exemple précédent (naissance ou disparition d'un couple de points singuliers) épuise toutes les bifurcations dans les familles génériques de champs de vecteurs sur une droite (cf. § 31). Dans le cas multidimensionnel, la naissance ou l'élimination d'un couple de points singuliers est également un cas générique. Que devient le portrait de phases dans ce cas?

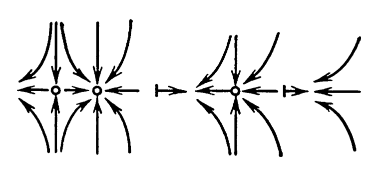


Fig. 124

Il s'avère que dans le cas d'une valeur caractéristique nulle la déformation topologiquement verselle d'un point singulier dégénéré générique dans $\mathbb{R}^n$ se déduit à partir de l'équation de l'exemple précédent par une simple suspension :

$$\begin{cases} \dot{x} = \pm x^2 + \varepsilon, & x \in \mathbb{R}, \ \varepsilon \in \mathbb{R}, \\ \dot{y} = -y, & y \in \mathbb{R}^{n_-}, \\ \dot{z} = z, & z \in \mathbb{R}^{n_+}, \end{cases} \tag{1}$$

où n_- et n_+ sont les nombres de racines de l'équation caractéristique dans le demi-plan de gauche et de droite respectivement. Pour $n = 2$, par exemple, ce système décrit la fusion d'un nœud et d'un col (fig. 124). Pour $\varepsilon = 0$ on obtient un col-nœud.

Dans le § 31 nous avons appelé points de bifurcation les points du produit direct de l'espace des phases par l'espace des valeurs du paramètre pour lesquelles l'équation caractéristique possède une racine nulle.

T h é o r è m e. *Les familles génériques qui, au voisinage de chaque point de bifurcation, sont topologiquement équivalentes à la famille* (1) *au voisinage de l'origine des coordonnées, forment un ensemble partout*

*dense *) dans l'espace des familles à un paramètre de champs de vecteurs.*

Il est commode de démontrer ce théorème en le réduisant au cas $n = 1$ pour lequel ce théorème est évident (et a été démontré plus haut). Cette réduction qui permet de diminuer la dimension de l'espace des phases au minimum peut être effectuée une fois pour toutes dans la situation la plus générale.

Considérons une famille locale de champs de vecteurs, dépendant d'un paramètre de dimension finie $(v;\, x_0,\, \varepsilon_0)$. Pour condenser l'écriture on admettra que $x_0 = 0 \in \mathbb{R}^n$, $\varepsilon_0 = 0 \in \mathbb{R}^k$.

Supposons que le champ $v(\cdot, 0)$ présente un point singulier en $x = 0$ et que l'équation caractéristique correspondante possède n_- (respectivement n_+, n_0) racines dans le demi-plan de gauche (respectivement dans le demi-plan de droite, sur l'axe imaginaire).

T h é o r è m e. *Sous les conditions admises la famille est topologiquement équivalente à la suspension à une famille d'espace des phases à n_0 dimensions*

$$\begin{cases} \dot{\xi} = w(\xi, \varepsilon), & \xi \in \mathbb{R}^{n_0}, \ \varepsilon \in \mathbb{R}^k, \\ \dot{y} = -y, & y \in \mathbb{R}^{n_-}, \\ \dot{z} = z, & z \in \mathbb{R}^{n_+}. \end{cases}$$

◀ On trouvera la démonstration de ce théorème dans l'article de A. C h o c h i t a ï c h v i l i, Bifurcations de type topologique d'un champ de vecteurs au voisinage d'un point singulier, Travaux du séminaire de I. Petrovski, 1, 1975, pages 279-309, en russe (voir également « Analyse fonctionnelle et ses applications », 6, 2 (1972), 97-98, où ce théorème a été formulé pour la première fois).

La démonstration est effectuée selon le même schéma que pour le théorème d'Anossov des C-systèmes: le plus gros de la démonstration consiste à construire cinq feuilletages (un contractant, un dilatant, un neutre, un non contractant, un non dilatant) dans le produit direct de l'espace des phases par l'espace des paramètres. [Les paramètres peuvent être traités comme des variables supplémentaires auxquelles est associée l'équation $\dot{\varepsilon} = 0$, mais il faut veiller à ce que les plans $\varepsilon = \text{const}$ se transforment en des plans analogues par les changements considérés.] ▶

Indépendamment des besoins de la théorie des bifurcations, l'existence des cinq feuilletages a été prouvée par E. Tikhonova (E. T i k h o n o v a, Analogie et homomorphisme de systèmes perturbés et non perturbés de matrice triangulaire partitionnée, Equations différentielles, 6, 7 (1970), 1221-1229, et par M. Hirsch,

---

*) Comme toujours l'ensemble des familles génériques est l'intersection d'un nombre dénombrable d'ensembles ouverts. Cet ensemble est ouvert si le domaine de définition des familles est compact, ou si la topologie utilisée est fine.

C. Pugh et M. Shub (M. W. Hirsch, C. C. Pugh, M. Shub, Invariant manifolds, BAMS, 76, 5 (1970), 1015-1019). Le cas $n_+ = 0$ a été étudié antérieurement par Pliss (V. Pliss, Principe de réduction en théorie de stabilité du mouvement, Izv. AN SSSR, Ser mathem., 28, 6 (1964), 1297-1324).

L'équation différentielle du système réduit $\dot{\xi} = w(\xi, \varepsilon)$ est réalisée dans le système initial sur une *sous-variété neutre* différentiable et dépendant différentiablement de ε, de dimension n_0 dans l'espace des phases. La sous-variété neutre est différentiable un nombre fini de fois (ce nombre croît pour $\varepsilon \to 0$) et est définie de façon non unique (comme le montrent des exemples simples).

Néanmoins le comportement des trajectoires y compris au point de bifurcation est déterminé pour l'équation complète par ce qui se passe sur la sous-variété neutre indiquée (et en particulier ne dépend pas du choix de la variété neutre).

A. Chochitaïchvili a prouvé également que la versalité de la déformation initiale est équivalente à celle de la déformation réduite (c'est-à-dire à celle de la déformation initiale sur la variété neutre).

Donc, l'étude topologique des singularités locales des portraits de phases au voisinage de points singuliers, bifurcations comprises, peut être limitée au cas où toutes les racines de l'équation caractéristique sont imaginaires pures. Le passage au cas général s'effectue par une simple suspension (par un produit direct par le col standard $\dot{y} = -y$, $\dot{z} = z$).

Exemple. De ce qui précède il suit en particulier qu'une trajectoire (et une seule) (pour les valeurs des paramètres proches de la valeur de bifurcation) joint un point à l'autre du couple de points singuliers né dans une famille générique à un paramètre de champs de vecteurs.

§ 33. Perte de stabilité de la position d'équilibre

On étudie ici les bifurcations du portrait de phases d'une équation différentielle lorsqu'un couple de racines de l'équation caractéristique devient imaginaire.

A. Exemple: Pertes douce et brusque de stabilité.

Commençons par un exemple de familles à un paramètre de champs de vecteurs sur le plan remontant à Poincaré et Andronov. Mettons-le sous la forme complexe

$$\dot{z} = z(i\omega + \varepsilon + cz\bar{z}), \qquad (1)$$

où $z = x + iy$ est une coordonnée complexe sur le plan $\mathbb{R}^2$ traité comme le plan de la variable complexe z.

Dans la formule précédente, ω et c sont des constantes non nulles réelles que l'on peut si l'on veut prendre égales à ± 1 ; ε est un paramètre réel.

Le point $z = 0$ est un point d'équilibre de type foyer pour tout ε. Ce foyer est stable pour $\varepsilon < 0$ et instable pour $\varepsilon > 0$. Pour $\varepsilon = 0$ l'approximation linéaire nous donne un centre qui est stable ou instable selon que $c < 0$ ou $c > 0$.

Lorsqu'on étudie localement les points singuliers, on remarque qu'à l'instant où $\varepsilon = 0$ le point singulier perd sa stabilité, mais on omet une circonstance importante liée à cette perte de stabilité: la naissance d'un cycle limite (cf. fig. 127). Pour éviter cette erreur il faut considérer le voisinage du point 0 dans le (z, ε) -espace et non dans le z-espace pour ε fixe.

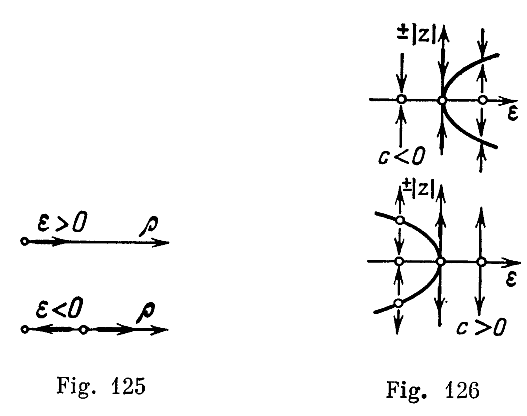


Fig. 125

Fig. 126

Il est commode d'étudier le voisinage du 0 dans le (z, ε)-espace de la manière suivante. Soit la fonction $\rho(z) = z\bar{z}$. De (1) on trouve l'équation pour ρ:

$$\dot{\rho} = 2\rho(\varepsilon + c\rho), \quad \rho \geqslant 0.$$

Il est facile d'étudier cette famille d'équations sur le rayon $\rho \geqslant 0$. Il existe deux points singuliers: le point singulier $\rho = 0$, qui apparaît pour chaque ε, et le point $\rho = -\varepsilon/c$ (si ε et c sont de signes contraires). Pour $c > 0$ le champ de vecteurs est de l'une des formes indiquées sur la figure 125 selon le signe de ε.

Au point $\rho = 0$ du plan z correspond l'origine des coordonnées, au point $\rho = -\varepsilon/c$ un cycle limite (réel uniquement lorsque ε et c sont de signes contraires).

Pour mieux comprendre la situation on portera ε sur un axe, et $|z|$, sur un autre de part en part. Le comportement du cycle lorsque le paramètre varie est alors illustré selon le signe de c par l'un des deux diagrammes de la figure 126. Le rayon du cycle est donc proportionnel à $\sqrt{|\varepsilon|}$.

Traitons d'abord le cas $c < 0$. Lorsque ε passe par 0, le foyer de l'origine des coordonnées perd sa stabilité. Pour $\varepsilon = 0$, l'origine des coordonnées est également un foyer stable mais non structurellement stable : les trajectoires ne se rapprochent pas exponentiellement de 0 (fig. 127).

Pour $\varepsilon > 0$ les trajectoires s'éloignent du foyer à une distance proportionnelle à $\sqrt{\varepsilon}$ et s'enroulent autour du cycle limite stable. Donc, lorsque ε passe par 0, $c < 0$, la perte de stabilité s'accompagne de la naissance d'un cycle limite stable dont le rayon croît comme $\sqrt{\varepsilon}$.

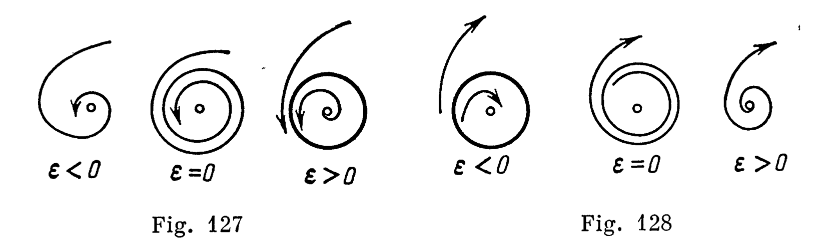


Fig. 127 Fig. 128

En d'autres termes, l'état stationnaire perd sa stabilité et il apparaît un régime périodique stable dont l'amplitude est proportionnelle à la racine carrée de l'écart du paramètre par rapport à la valeur critique. Les physiciens parlent dans ce cas d'une *excitation douce d'auto-oscillations.*

Lorsque $c > 0$ (fig. 128), on a un cycle limite instable pour $\varepsilon < 0$. Lorsque ε tend vers 0, le cycle prend une position d'équilibre qui pour $\varepsilon < 0$ était un foyer stable. Pour $\varepsilon = 0$ le foyer devient instable (l'instabilité est faible, non exponentielle). Pour ε positif le foyer est instable déjà en approximation linéaire.

Ce cas de perte de stabilité est appelé *excitation brusque* pour la raison suivante.

Imaginons-nous qu'un système se trouve à l'état d'équilibre au voisinage d'une position d'équilibre stable et que cette position perde sa stabilité lorsque le paramètre varie. Pour $c > 0$, lorsque ε tend vers 0 par valeurs négatives (ou même bien avant), les perturbations qui sont toujours présentes repoussent le système du voisinage de la position d'équilibre et ce système saute d'un coup sur un autre régime (par exemple sur une position d'équilibre lointaine, un cycle limite ou un ensemble attractif plus compliqué). Donc, lorsque le paramètre varie de façon continue, le régime du mouvement varie par saut, brusquement.

Pour $c < 0$ l'amplitude des auto-oscillations engendrées ne dépend pas différentiablement (singularité radicale) du paramètre mais continûment ; dans ce sens le régime du mouvement varie de façon continue, doucement.

En étudiant l'équation (1) nous avons essentiellement utilisé le point de vue « versel »: si au lieu d'un voisinage dans le (z, ε)-espace nous avions considéré un voisinage dans le z-espace à ε fixe, nous aurions omis les cycles limites. Cela concorde avec le fait qu'une singularité de codimension k doit être étudiée dans une famille k-paramétrique: notre cas de codimension 1 est compris dans une famille à un paramètre.

L'exemple considéré épuise effectivement les bifurcations du portrait de phases dans les familles à un paramètre génériques, qui apparaissent lorsque la position d'équilibre perd sa stabilité et, d'une façon plus générale, lorsqu'un couple de racines de l'équation caractéristique traverse l'axe imaginaire.

B. Théorème de Poincaré-Andronov

Considérons une famille à un paramètre de champs de vecteurs.

Supposons que lorsque le paramètre prend la valeur 0, le champ présente un point singulier 0 tel que les racines de l'équation caractéristique soient imaginaires pures (l'espace des phases est de dimension 2).

T h é o r è m e. *Chaque famille locale générique (c'est-à-dire générique parmi les familles possédant les propriétés indiquées) est topologiquement équivalente à la famille de l'exemple précédent.*

◀ On se servira de la méthode de Poincaré pour réduire cette équation à sa forme normale. Lorsque le paramètre prend la valeur 0, on a une résonance qui ne se manifeste pas pour les valeurs voisines non nulles du paramètre. Les termes résonnants correspondants ne peuvent pas être éliminés pour la valeur 0 du paramètre et peuvent l'être pour les valeurs voisines. Si pour des valeurs non résonnantes proches de zéro du paramètre nous allons éliminer les termes devenus résonnants pour la valeur zéro du paramètre, notre substitution dépendra discontinûment du paramètre et le rayon du voisinage dans lequel nous étudions le portrait de phases tendra vers 0 lorsque le paramètre tendra vers la valeur résonnante.

C'est pourquoi nous n'éliminerons pas les termes devenant résonnants pour $\varepsilon = 0$ non seulement pour la valeur 0 du paramètre, mais pour toutes les valeurs proches. On obtient en définitive une substitution dépendant différentiablement du paramètre par laquelle le système ne contiendra que des termes devenant résonnants pour la valeur 0 du paramètre et un résidu d'ordre supérieur aussi grand que l'on veut par rapport à la distance au point singulier. Nous nous proposons d'étudier la bifurcation dans la famille obtenue en négligeant le reste sous réserve qu'il n'influe pas sur la topologie des modifications (sinon nous en tiendrons compte).

Le programme défini est général pour de nombreux problèmes sur les bifurcations. Voyons ce qu'il donne dans notre cas concret de passage d'un couple de racines de l'équation caractéristique par

17–0743

l'axe imaginaire. La résonance est de la forme $\lambda_1 + \lambda_2 = 0$ ($\lambda_{1,2} = \pm i\omega$). Donc, dans les coordonnées propres du plan complexifié $\mathbb{C}^2$, la forme normale s'écrit (cf. § 23)

$$\dot{z}_1 = \lambda_1(\varepsilon) z_1 + a_1(\varepsilon) z_1^2 z_2 + \ldots,$$
$$\dot{z}_2 = \lambda_2(\varepsilon) z_2 + b_1(\varepsilon) z_1 z_2^2 + \ldots$$

On remarquera que, l'équation initiale étant réelle, la base propre peut être constituée de vecteurs conjugués complexes et les substitutions normalisantes, choisies réelles. Dans ce cas la deuxième équation est conjuguée de la première. Par ailleurs dans le plan réel $z_2 = \bar{z}_1$, par suite, on peut écrire une première équation en y remplaçant z_1 par z, et z_2 par $\bar{z}$. On peut considérer que cette équation représente le système initial dans le plan réel $\mathbb{R}^2$ sous la forme d'une équation (non holomorphe) sur la droite complexe $\mathbb{C}^1$ avec la coordonnée z

$$\dot{z} = \lambda_1(\varepsilon) z + a_1(\varepsilon) z^2\bar{z} + \ldots;$$

les points de suspension représentent un résidu d'ordre 5 par rapport à $|z|$.

Nous sommes donc conduits à l'étude de la famille

$$\dot{z} = \lambda_1(\varepsilon) z + a_1(\varepsilon) z^2\bar{z}.$$

On fera cette étude comme pour l'exemple spécial examiné au point A. Voici la correspondance entre les notations:

exemple du point A	$i\omega$	ε	c
famille générique	$\lambda_1(0)$	$\mathrm{Re}\,\lambda_1(\varepsilon)$	$\mathrm{Re}\,a_1(0)$

Dans les familles génériques on a

$$\lambda_1(0) \neq 0, \quad d/d\varepsilon\, \mathrm{Re}\,\lambda_1(\varepsilon)|_{\varepsilon=0} \neq 0, \quad \mathrm{Re}\,a_1(0) \neq 0.$$

La bifurcation consiste en la naissance ou la disparition du cycle limite (naissance pour $\mathrm{Re}\,a_1(0)$ et $\mathrm{Re}\,d\lambda_1/d\varepsilon\,|_0$ de signes contraires).

Si les trois grandeurs indiquées plus haut, qui sont généralement non nulles, sont différentes de zéro, alors la considération du résidu négligé ne modifie pas les bifurcations. Ce qu'on prouve aisément en considérant la dérivée de la fonction $\rho = |z|^2$ le long de notre champ de vecteurs:

$$\dot{\rho} = 2\rho(\mathrm{Re}\,\lambda_1(\varepsilon) + \rho\,\mathrm{Re}\,a_1(\varepsilon) + O(\rho^2)).$$

On voit avec cette formule que $O(\rho^2)$ n'influe pas sur la bifurcation du portrait de phases dans un voisinage (ne dépendant pas de ε) de l'origine des coordonnées. ▶

Le théorème précédent en fait était connu de Poincaré ; l'énoncé et la démonstration sont dus à A. Andronov (A. Andronov, Application de la théorie de Poincaré des « points de bifurcation » et « changement de stabilité » à des systèmes auto-oscillants simples, C. R. Ac. Sci Paris, 189, 15 (1929), 559-561 ; A. Andronov, E. Leontovitch-Andronova, « Certains cas de dépendance des mouvements périodiques par rapport à un paramètre. Outch. zapiski, GGU, 1939, Vyp. 6, page 3 (Œuvre de A. Andronov, 189-216)). R. Thom, à qui j'ai enseigné cette théorie en 1965, l'a faite largement connaître sous le nom de « bifurcation d'E. Hopf » (cf., par exemple, l'ouvrage de S. Smale et M. Hirsch *)).

C. Cas multidimensionnel.

En combinant le théorème de Poincaré-Andronov et le théorème de réduction (§ 32) on est conduit au

Théorème. *La déformation verselle topologique d'un point singulier d'un champ de vecteurs générique de $\mathbb{R}^n$ dont l'équation caractéristique possède un couple de racines imaginaires pures se déduit par une simple suspension à partir du système de Poincaré-Andronov*

$$\begin{cases} \dot{z} = z(i + \varepsilon \pm z\bar{z}), & z \in \mathbb{C}^1, \quad \varepsilon \in \mathbb{R}; \\ \dot{u} = -u, & u \in \mathbb{R}^{n_-}, \\ \dot{v} = v, & v \in \mathbb{R}^{n_+}, \quad n = n_- + n_+ + 2. \end{cases}$$

L'étude de ces systèmes ne présente désormais aucune difficulté.

Exemple. Supposons que $n = 3$, $n_+ = 0$ et que $z\bar{z}$ est précédé du signe moins. Dans ce cas, le théorème affirme que dans un voisinage de l'origine de coordonnées, ne dépendant pas de ε, est engendré au passage d'un couple de valeurs propres par l'axe imaginaire un cylindre invariant de rayons $\sqrt{\varepsilon}$ attirant les trajectoires voisines. Sur le cylindre on a un cycle stable sur lequel toutes les trajectoires viennent s'enrouler à la longue. Donc, ce cas correspond à une perte douce de stabilité avec apparition d'auto-oscillations.

La singularité mentionnée a été étudiée par de nombreux auteurs. En particulier, E. Hopf a traité la naissance du cycle dans le cas multidimensionnel (E. Hopf, Abzweigung einer periodischen Lösung von einer stationären Lösung, Bereich, Sächs. Acad. Wiss. Leipzig, Math. Phys. Kl. 94, 19 (1942), 15-25). D'autres résultats ont été obtenus par Y. Neimark et N. Brouchlinskaïa.

Cependant le théorème général formulé plus haut et donnant une étude complète des bifurcations du portrait de phases (et pas seulement des bifurcations du cycle) n'a été démontré que dans le travail cité plus haut de A. Chochitaïchvili de réduction à l'aide des résultats en dimension 2 d'Andronov-Poincaré.

D. Application à la théorie de la stabilité hydrodynamique.

Les phénomènes examinés plus haut se rencontrent souvent dans des situations concrètes : les systèmes mécaniques, physiques, chimiques, biologiques et économiques perdent leur stabilité à chaque pas.

---

*) M. Hirsh, S. Smale, Differential equations, Dynamical systems and Linear Algebra, New York, A.P., 1974.

A titre d'exemple nous allons étudier ici un problème spécial de cette nature: la perte de stabilité d'un écoulement stationnaire d'un fluide visqueux incompressible.

Soient D un domaine rempli de liquide, v le champ de vitesses du liquide. L'écoulement est décrit par les équations de Navier-Stokes

$$\frac{\partial v}{\partial t} + (v\nabla,\ v) = \nu\Delta v - \operatorname{grad} p + f, \quad \operatorname{div} v = 0,$$

où le coefficient ν représente la viscosité, f le champ des forces massiques non potentielles; la pression p se détermine à partir de la condition d'incompressibilité. Sur la frontière du domaine D on pose, par exemple, des conditions d'adhérence ($v|_{\partial D} = 0$).

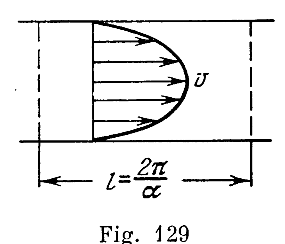


Fig. 129

On suppose que le champ initial des vitesses définit le mouvement ultérieur de telle sorte que l'équation définisse un système dynamique dans l'espace de dimension infinie des champs de vecteurs de divergence nulle, nuls sur la frontière du domaine D. Ceci n'a en fait été démontré que dans le cas bidimensionnel. Les problèmes d'existence, d'unicité et les propriétés des solutions des équations de Navier-Stokes ont fait l'objet d'une multitude de travaux, cependant, les principaux d'entre eux restent entiers.

Etudions par exemple l'écoulement de Poiseuille (de profil parabolique des vitesses, fig. 129) dans un canal plan. L'écoulement de Poiseuille est un point stationnaire dans notre système dynamique dans l'espace fonctionnel quelle que soit la valeur ν de la viscosité. Cette position d'équilibre est stable pour une viscosité assez grande, cependant elle devient instable lorsque la viscosité décroît. Nous pouvons étudier ce qu'il adviendra à l'aide du théorème du point C.

Il faut certes prendre des précautions particulières en raison de la dimension infinie du problème. Il faut espérer que cette dimension infinie n'est pas très dangereuse, car la viscosité amortit rapidement les harmoniques élevés de sorte que le système se transforme en fait en système de dimension finie pour une valeur non nulle quelconque du coefficient de viscosité. L'autre difficulté et que nous ne sommes pas sûrs que notre système est réellement générique: il faut vérifier cela par des calculs. Il semble naturel que le système de Navier-Stokes soit un système générique dans un domaine « générique » pour des forces massiques génériques, cependant l'écoulement de Poiseuille est assez spécial, ici il existe un grand groupe de symétries.

Bornons-nous à des perturbations dont le champ de vitesses le long du flot est périodique et de période l. Pour normer la vitesse de l'écoulement principal, on modifiera les forces extérieures proportionnellement à la viscosité de telle sorte que le débit Q du liquide soit constant ($f = \text{const}\cdot Q\nu$). Nous obtenons ainsi une famille

à deux paramètres l et ν. Généralement pour paramètre on considère les grandeurs inverses $\alpha = 2\pi/l$ (le nombre d'ondes), $\mathbf{R} = \text{const}\cdot Q/\nu$ (nombre de Reynolds). Donc, à une diminution de la viscosité entraînée par une instabilité correspond un accroissement du nombre de Reynolds.

Les calculs (qui sont pratiquement impossibles sans ordinateur) montrent que lorsque le nombre de Reynolds croît et pour une valeur critique de ce nombre, $\mathbf{R}_0 = \mathbf{R}_0(\alpha)$, un couple de racines complexes

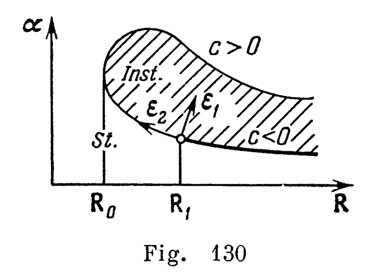


Fig. 130

traverse l'axe imaginaire du demi-plan stable dans le demi-plan instable. Donc nous avons affaire ici au cas de perte de stabilité donnant naissance ou détruisant un cycle limite.

Le signe du coefficient c qui indique si l'excitation des oscillations est brusque ou douce est également calculé. Pour décrire les résultats il est commode de représenter la frontière de stabilité sur le plan $(\alpha, \mathbf{R})$. Cette frontière a la forme d'une « langue » représentée sur la figure 130. Le point extrême gauche de cette « langue » est très important. Sa $\mathbf{R}$-coordonnée correspond à la première perte de stabilité, sa α-coordonnée définit la plus dangereuse longueur d'onde du point de vue de l'instabilité.

Il s'avère que le coefficient c est positif pour toute la région gauche et toute la région supérieure à la frontière de la stabilité, c'est-à-dire on a une excitation *brusque*. Donc, avant que le nombre de Reynolds passe par la valeur critique $\mathbf{R}_0$, quelque part dans l'espace des phases, loin du point stationnaire (c'est-à-dire loin de l'écoulement de Poiseuille), apparaît un régime oscillant *) sur lequel les petites perturbations repoussent le système lorsque le nombre de Reynolds tend vers $\mathbf{R}_0$. Ce nouveau régime est susceptible d'être un point stationnaire stable (ou, en terme d'hydrodynamique, un écoulement stationnaire distinct de l'écoulement de Poiseuille) ou un cycle limite (en terme d'hydrodynamique, un écoulement périodique), mais peut posséder une structure plus complexe, par

---

*) Ou bien une dérive à l'infini; mais ce ne sera probablement pas le cas ici, puisque à l'infini la vitesse est orientée vers l'origine des coordonnées en raison de l'amortissement dû à la viscosité.

exemple, peut être un mouvement quasi périodique sur un tore. Bien plus le régime engendré par l'excitation brusque peut être un C-système ou un système à caractère hyperbolique, c'est-à-dire un ensemble attractif avec des trajectoires non régulières, instables. Le spectre du système dynamique correspondant peut être continu malgré la finitude du nombre de degrés de liberté (c'est-à-dire la finitude de la dimension de l'ensemble attractif). Les expérimentateurs auraient appelé un tel régime écoulement turbulent.

Lorenz a publié en 1963 un travail (E. N. Lorenz, Deterministic nonperiodic flow, J. Atmos. Sci. **20** (1963), 130-141), dans lequel il étudie, le premier, un régime attractif non trivial dans un système d'espace des phases à trois dimensions, simulant la théorie hydrodynamique de convection.

Le système de Lorenz s'écrit

$$\dot{x} = -\sigma x + \sigma y, \quad \dot{y} = -xz + rz - y, \quad \dot{z} = xy - bz; \quad \sigma = 10, \quad r = 28, \quad b = 8/3.$$

Il semble que tous les modèles dans lesquels on a réussi jusqu'à maintenant à trouver des ensembles hyperboliques attractifs contiennent des membres de type pompage ou viscosité négative, inexistants dans les équations de Navier-Stokes. En tous les cas lorsqu'en 1964 j'ai essayé de trouver un ensemble attractif hyperbolique dans un espace des phases à six dimensions d'une approximation de Galerkine de l'équation de Navier-Stokes sur un tore à deux dimensions, avec une force extérieure sinusoïdale (en utilisant un programme de N. Vvédenskaïa), l'ensemble attractif était probablement un tore à trois dimensions (en raison peut-être d'un nombre de Reynolds trop petit). A ma connaissance on n'a pas encore trouvé les ensembles attractifs hyperboliques pour les équations de Navier-Stokes ou leurs approximations de Galerkine. Par ailleurs, l'expérience numérique décrite plus haut a inspiré de nombreux travaux sur l'application à l'hydrodynamique des flots géodésiques sur des groupes de difféomorphismes (V. Arnold, Sur la géométric différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. Ann. Inst. Fourier, Grenoble, 16, 1 (1966), 319-361; D. G. Ebin, J. Marsden, Groups of diffeomorphisms and the motion of an incompressible fluid, Ann. of Math. **92** (1970), 102-163).

M. Henon (M. Henon, A Two-dimensional Mapping with a Strange Attractor, Comm. Math Phys. **50** (1976), 69-77) a proposé un modèle très simple à trajectoires instables sur un ensemble attractif. Henon considère une « transformation de Cremona » quadratique sur le plan, de la forme $T = T_1 T_2 T_3$, où

$$T_1(x, y) = (y, x), \; T_2(x, y) = (bx, y), \; T_3(x, y) = (x, y + 1 - ax^2).$$

Il est intéressant de noter que dans l'expérience numérique effectuée pour $a = 1,4$, $b = 0,3$, l'attraction par un ensemble, possédant localement la forme d'un produit d'un ensemble de Cantor par un segment, est impossible à décrire dans le cadre des définitions existantes de l'hyperbolicité (il n'est pas exclu non plus que cet ensemble contienne les domaines d'attraction des cycles longs). Donc, les mathématiciens n'admettent pas que l'ensemble attractif de Henon est hyperbolique. Dans le même temps, du point de vue de l'expérimentateur, le mouvement du point représentatif sous l'action des itérations de la transformation T revêt manifestement un caractère turbulent stochastique (voici encore un exemple du danger présenté par le fétichisme des axiomes).

R. Plikine a construit des exemples d'ensembles attractifs vraiment hyperboliques sur le plan (R. Plikine, Sources et débits de A-difféomorphismes de surfaces, Matematitcheski zbornik, **94**, 2 (1974), 243-264). Plikine a construit

un difféomorphisme du domaine fermé comportant trois trous, représenté en haut de la figure 131, dans la partie hachurée du bas de la figure 131, possédant la propriété suivante: l'intersection des images du domaine par toutes les itérations du difféomorphisme est un ensemble attractif (la distance des images d'un point quelconque par les itérations à cet ensemble tend vers 0) ; cette intersection

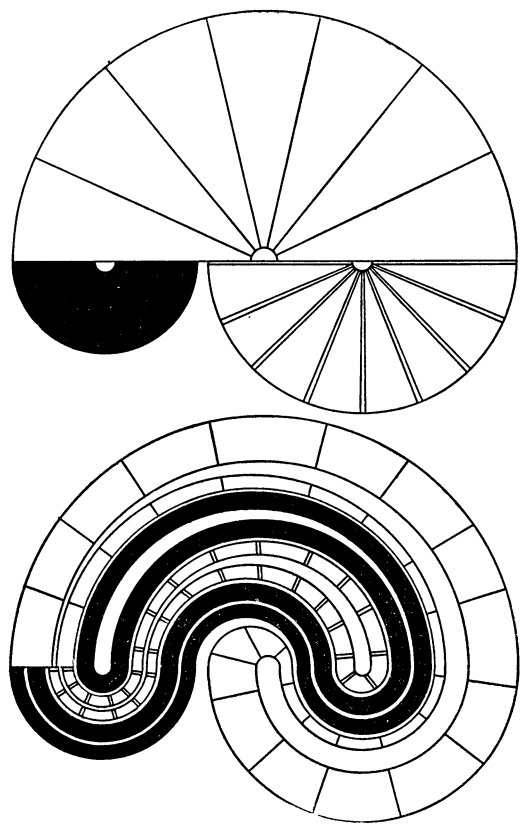


Fig. 131

localement est le produit d'un ensemble de Cantor par un segment, la distance entre les points proches sur chaque segment croît par les puissances du difféomorphisme.

Le livre de J. Marsden, M. McCracken, The Hopf Bifurcation and its Applications, Springer, 1976, contient une riche bibliographie des travaux portant sur la théorie des bifurcations et ses applications (plus de 350 titres).

La détermination du régime adopté par le système dynamique lors d'une perte de stabilité de l'écoulement de Poiseuille se trouve de l'avis des experts à la limite des possibilités des calculatrices modernes.

Dans cette situation il importe probablement de ne pas négliger les prévisions qualitatives que l'on peut faire sans calcul en s'appuyant sur la théorie générale des bifurcations exposée plus haut.

Le problème posé comporte deux paramètres, α et **R**. Donc, outre les singularités de codimension 1, on peut rencontrer des singularités de codimension 2. Etudions l'une d'entre elles, à savoir celle qui est liée au changement de signe de c. Les calculs montrent que, pour un nombre de Reynolds **R** assez grand, l'excitation brusque

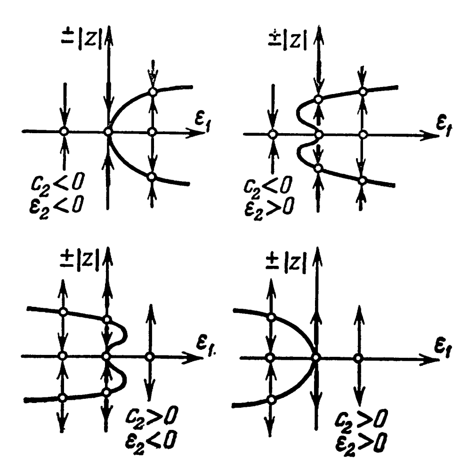


Fig. 132

à la partie inférieure de la frontière de perte de stabilité est remplacée par une douce. Pour comprendre ce qui se passe à cet instant, il faut construire une famille verselle à deux paramètres pour une telle dégénérescence double. Cette famille qui est facile à construire est de la forme

$$\dot{z} = z\,(i\omega + \varepsilon_1 + \varepsilon_2 z\bar{z} + c_2 z^2\bar{z}^2), \quad z \in \mathbb{C}$$

(les autres coordonnées de l'espace des phases, qui correspondent aux valeurs propres stables, n'ont pas été écrites). Le sens des paramètres ε_1 et ε_2 est clair de la figure 130; le caractère de la modification au point $\varepsilon_1 = \varepsilon_2 = 0$ est déterminé par le signe de c_2.

Si l'on pose $\rho = z\bar{z}$ comme plus haut, on obtient l'équation

$$\dot{\rho} = 2\rho\,(\varepsilon_1 + \varepsilon_2\rho + c_2\rho^2), \quad \rho \geqslant 0.$$

On distinguera les cas suivants selon le signe de ε et c:

1°. $c_2 < 0$, $\varepsilon_2 < 0$. Lorsque ε_1 change de signe et devient positif, le système arrive doucement à un régime auto-oscillant stable périodique (fig. 132).

2°. $c_2 < 0$, $\varepsilon_2 > 0$. Lorsque ε_1 change de signe et devient positif, le système arrive brusquement à un régime auto-oscillant périodique stable engendré avant la perte de stabilité de la position d'équilibre en même temps que le régime oscillant instable qui a atterri sur la position d'équilibre au moment de la perte de stabilité.

On aurait pu étudier le cycle limite stable indiqué plus haut au voisinage du point de changement du régime brusque en régime doux, puisqu'il est proche de la position d'équilibre. Cependant ce cycle peut être analytiquement prolongeable (au loin de la position d'équilibre) pour d'autres valeurs des paramètres (α, **R**) ; on voit qu'on peut chercher ce prolongement à l'aide du prolongement analytique du cycle instable qui a atterri sur la position d'équilibre lors de la perte brusque de stabilité. Le cycle stable indiqué est susceptible d'être le régime où arrive le système lors de la perte de stabilité du régime stationnaire.

3°. $c_2 > 0$, $\varepsilon_2 < 0$. La perte de stabilité est douce, mais le cycle limite disparaît rapidement en fusionnant avec un cycle instable venu de loin, après quoi un nouveau régime est excité brusquement.

4°. $c_2 > 0$, $\varepsilon_2 > 0$. On a une excitation brusque ordinaire.

Donc, quel que soit le signe de c_2, par un choix convenable du signe de ε_2, notre analyse nous permet d'établir un fait nouveau relativement à l'analyse « à un paramètre »: pour $c_2 < 0$ on trouve un régime final à excitation brusque, et pour $c_2 > 0$ on constate que le régime final à excitation douce est éphémère. Pour établir lequel de ces deux cas ($c_2 < 0$ ou $c_2 > 0$) a lieu en réalité il faut effectuer des calculs assez volumineux.

En théorie de la stabilité hydrodynamique on rencontre diverses singularités de la frontière de stabilité et du diagramme du décrément, de sorte que les résultats du § 30 peuvent trouver une application. Pour se servir de la théorie générale des bifurcations en théorie de la stabilité hydrodynamique, il serait important d'étudier les cas génériques dans les problèmes comportant divers groupes de symétrie, puisque dans de nombreux problèmes d'hydrodynamique, le domaine d'écoulement D est soumis à l'action de tel ou tel groupe de symétries (par exemple, le groupe des translations dans le problème d'écoulement de Poiseuille; les représentations de ce groupe interviennent dans l'étude sous la forme d'un paramètre α).

Le comportement du liquide après la perte de stabilité de l'écoulement stationnaire fait l'objet de nombreux travaux (voir par exemple L. Landau, E. Lifchits, Mécanique des milieux continus, M., Gostekhizdat, 1954 ; § 27). Dans ces travaux on suppose d'habitude un régime doux d'excitation des auto-oscillations et on étudie la perte de stabilité du cycle limite. Landau a supposé que seront engendrés des mouvements quasi périodiques avec un nombre grandissant de fréquences; il ne fait aucun doute que cela s'explique par le fait que les autres systèmes dynamiques lui étaient inconnus.

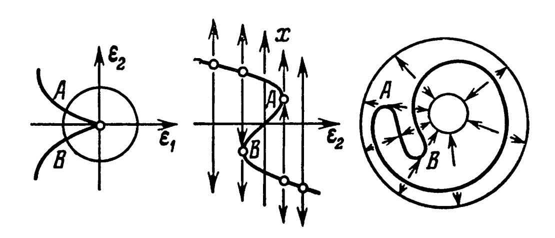
En 1965, j'ai présenté la théorie exposée plus haut à l'occasion d'un séminaire de Thom à l'Institut de recherches supérieures à Bure. D. Ruelle, F. Takens (D. Ruelle, F. Takens, On the nature of turbulence, Comm. Math. Phys. **20** (1971), 167-192; **23** (1971)) ont construit 5 années plus tard des exemples de perte de stabilité d'un cycle avec apparition d'un régime plus compliqué qu'un mouvement quasi périodique; cependant leur exemple revêt un caractère

exotique, car il correspond à un sous-espace de l'espace des paramètres de déformation très maigre métriquement (bien qu'ouvert). J. B. McLaughlin, P. C. Martin passent en revue les travaux empiriques dans leur ouvrage « Transition to turbulence of a statically stressed fluid system », Phys. Rev Letters, 33 (1974); Phys. Rev. A 12 (1975), 186-203.

A noter que l'applicabilité des résultats des travaux indiqués implique une perte de stabilité douce, tandis que ce régime est brusque pour l'écoulement de Poiseuille.

E. Singularités de codimension 2.

Les cas examinés plus haut (naissance et disparition d'un couple de points singuliers, naissance ou disparition d'un cycle limite d'un



Fiig. 133

point singulier) épuisent les bifurcations des portraits de phases au voisinage d'un point singulier pour les familles génériques à un paramètre de champs de vecteurs.

Dans les familles à deux paramètres ces singularités se rencontrent sur des lignes du plan des paramètres, mais à part elles on trouvera des singularités plus complexes (en certains points du plan des paramètres). Parmi ces singularités les cinq suivantes sont inévitables par une petite déformation de la famille à deux paramètres.

1°. *Une racine nulle avec une dégénérescence supplémentaire.* Exemple:

$$\dot{x} = \pm x^3 + \varepsilon_1 x + \varepsilon_2, \quad x \in \mathbb{R}$$

(fig. 133). Il est aisé de vérifier que cette déformation est (topologiquement) verselle; dans le cas multidimensionnel la déformation verselle s'obtient par une suspension d'un col.

Le diagramme de bifurcation (pour le cas $+x^3$) est représenté à gauche de la figure 133. La parabole semi-cubique partage le plan des paramètres en deux régions. Dans la plus petite région le système possède trois positions d'équilibre au voisinage de $x = 0$, dans la plus grande, une. Les modifications du portrait de phases, lorsque le paramètre contourne le point $\varepsilon = 0$ le long d'un petit cercle, sont représentées sur la figure 133 à droite. Le produit direct de ce cercle par l'espace des phases (à une dimension) est un anneau circulaire sur lequel les positions d'équilibre forment une courbe fermée,

quant au comportement des vecteurs du champ, il est clair à partir de la figure 133.

2°. *Un couple imaginaire avec une dégénérescence supplémentaire.* Exemple :

$$\dot{z} = z\,(i\omega + \varepsilon_1 + \varepsilon_2 z\bar{z} \pm z^2\bar{z}^2), \qquad z \in \mathbb{C}.$$

Le diagramme des bifurcations est composé de la droite $\varepsilon_1 = 0$ et d'une demi-parabole tangente en 0 à cette droite ; il est représenté sur la figure 134 pour le cas où $z^2\bar{z}^2$ est précédé du signe « $+$ ».

Les modifications du portrait de phase, lorsque le paramètre contourne le point 0 le long d'un petit cercle, sont représentées sur

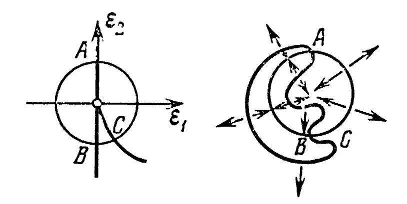


Fig. 134

la figure 134 à droite. L'anneau est le produit direct d'un cercle du plan des paramètres par la droite sur laquelle sont portées $\pm\ |z|$. Le cercle correspond à la position d'équilibre $z = 0$, chaque cycle limite est représenté par les deux points d'intersection d'un rayon avec la ligne $\varepsilon_1 + \varepsilon_2\,|z|^2 + |z|^4 = 0$.

Lorsque $z^2\bar{z}^2$ est précédé du signe « — », le diagramme des bifurcations et la famille sur le cercle sont analogues.

3°. *Deux couples imaginaires.*

4°. *Un couple imaginaire et une racine nulle.*

L'étude de ces cas n'a pas été encore poussée à un point tel que l'on puisse écrire les familles verselles ; bien plus on ne sait pas s'il existe dans le cas de 2 couples imaginaires une famille à deux paramètres (ou à la rigueur à un nombre fini de paramètres) topologiquement verselle (même sous l'hypothèse d'une incommensurabilité normale du rapport des fréquences lors de leur passage simultané d'un demi-plan dans l'autre).

Et enfin, le dernier cas de codimension 2 :

5°. *Deux racines nulles.* Exemple : la famille d'équations sur le plan

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = \varepsilon_1 + \varepsilon_2 x_1 + x_1^2 \pm x_1 x_2 \end{cases}$$

avec les paramètres (ε_1, ε_2). Le diagramme des bifurcations partage le plan ε en 4 régions notées A, B, C, D sur la figure 135, qui illustre le cas où x_1x_2 est précédé du signe « + ».

Les portraits de phases correspondant à chacune de ces 4 régions du plan ε sont représentés sur la figure 135. Aux branches du diagramme des bifurcations correspondent les systèmes à singularités de codimension 1, représentés sur la figure 135 (P, Q, R, S).

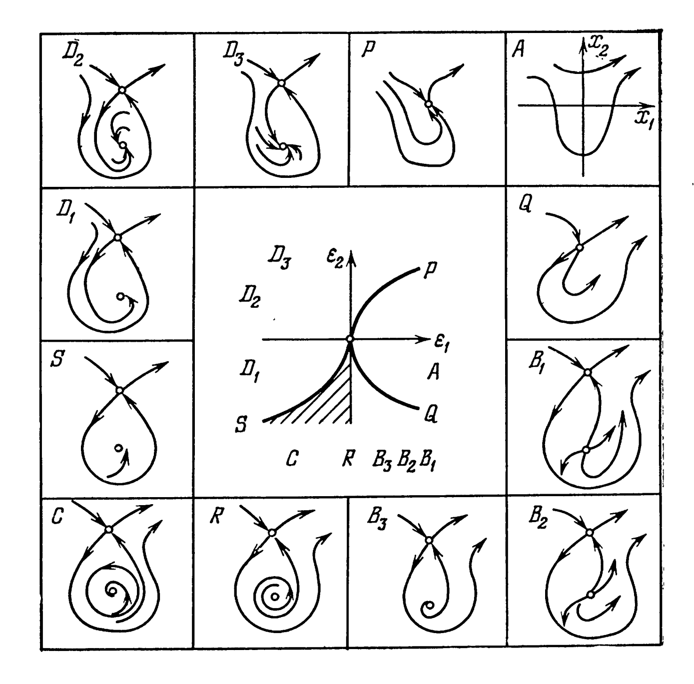


Fig. 135

On remarquera que la bifurcation sur la branche S (la naissance d'un cycle à partir de la boucle de la séparatrice) ne figure pas dans notre classification des singularités de codimension 1, car elle n'est pas locale (au voisinage du point singulier), mais globale. Nous voyons donc que lorsque le nombre de paramètres de la famille croît, les bifurcations globales de codimensions inférieures commencent à jouer un rôle dans l'étude locale des bifurcations des points singuliers. D'où il résulte que lorsque le nombre des paramètres est suffisant, on se heurte dans le problème local aux mêmes difficultés de non-densité partout des systèmes structurellement stables, mises en évidence par Smale dans le problème global des champs de vecteurs sur une variété (cf. § 15).

Si x_1x_2 est précédé du signe « — », les bifurcations se ramènent aux précédentes par changement des signes de t et x_2.

Théorème. *Les champs de vecteurs génériques dont l'équation caractéristique admet deux racines nulles possèdent au point singulier du plan des phases une déformation topologiquement verselle à deux paramètres équivalente à l'une des deux déformations considérées plus haut.*

Autrement dit, *une famille générique à deux paramètres d'équations différentielles sur le plan ayant pour une certaine valeur du paramètre un point singulier, dont l'équation caractéristique possède deux racines nulles, se ramène à la forme indiquée plus haut par un changement continu des paramètres et un changement continu des coordonnées dans l'espace des phases, dépendant continûment des paramètres.*

Ce théorème qui a été démontré par R. Bogdanov en 1971 a été publié pour la première fois dans l'aperçu de V. Arnold, Cours sur les bifurcations et les familles verselles, UMN, **27**, 5 (1972), 119-184 (en russe). Takens a annoncé un résultat analogue en 1974. La démonstration de la versalité n'est pas facile; la difficulté essentielle réside dans l'étude de l'unicité du cercle limite. R. Bodganov tourne cette difficulté par des raisonnements non triviaux sur le comportement des intégrales elliptiques en fonction du paramètre. Voir R. Bogdanov, Bifurcations du cycle limite d'une famille de champs de vecteurs sur le plan, Travaux du séminaire de I. Petrovski, 1976, Vyp. 2, 23-36 (en russe); Déformation verselle d'un point singulier d'un champ de vecteurs sur le plan dans le cas de valeurs propres nulles, idem, 37-65 (en russe).

§ 34. Perte de stabilité des auto-oscillations

Le problème suivant (par ordre de difficulté après la modification des portraits de phases au voisinage des positions d'équilibre) est le problème des modifications d'une famille de trajectoires au voisinage d'une trajectoire fermée. Ce problème n'a pas encore été entièrement résolu et probablement est irrésoluble dans un certain sens. Néanmoins les méthodes générales de la théorie de bifurcations permettent de recueillir une information importante sur ces modifications; ce paragraphe passe brièvement en revue les principaux résultats obtenus dans ce domaine.

A. Monodromie et multiplicateurs.

Considérons une trajectoire fermée d'un système d'équations différentielles. Etudions les modifications du comportement des trajectoires au voisinage de la trajectoire donnée par une petite déformation de l'équation.

Les décompositions en trajectoires au voisinage de la trajectoire fermée générique (à un homéomorphisme près du voisinage) sont en nombre fini. Pour les décrire choisissons un point O sur la trajectoire fermée. Menons par ce point une petite surface (de codimension 1 dans l'espace des phases) transversale à la trajectoire fermée. Les trajectoires issues des points de cette surface suffisamment

proches du point O recoupent cette surface après avoir contourné la trajectoire fermée. On obtient ainsi une application d'un voisinage du point O de la surface transversale dans elle-même. Cette application s'appelle *application de Poincaré* (fig. 136).

Le point O est un point fixe de l'application. Considérons la linéarisée de l'application de Poincaré au point O. Cet opérateur linéaire s'appelle *opérateur de monodromie.*

Les valeurs propres de l'opérateur de monodromie sont dites *multiplicateurs* de la trajectoire fermée initiale. L'opérateur de

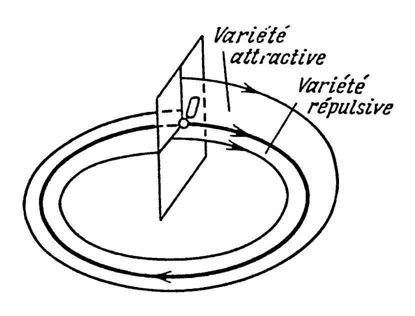


Fig. 136

monodromie est donnée par les solutions d'une équation linéaire à coefficients périodiques (une équation aux variations normales le long de la trajectoire fermée).

Supposons que les modules de tous les multiplicateurs sont inférieurs à l'unité. On démontre alors que les prolongements en avant de toutes les trajectoires voisines sont attirés par la trajectoire fermée. Si le module d'au moins un multiplicateur est plus grand que l'unité, il existe des trajectoires qui s'éloignent de la trajectoire fermée (autrement dit, se rapprochent d'elle pour $t \to -\infty$).

Dans le cas général, quelques valeurs propres sont situées à l'intérieur du cercle unitaire, et quelques valeurs, à l'extérieur. Dans ce cas les trajectoires attirées par la trajectoire fermée forment, ce qu'on peut prouver, une *variété attractive* dont l'intersection avec notre surface transversale possède une dimension égale au nombre de multiplicateurs situés à l'intérieur du cercle unité. De façon analogue, les trajectoires asymptotiques à la trajectoire fermée pour $t \to -\infty$ forment une *variété répulsive.* La dimension de son intersection avec la surface transversale est égale au nombre de multiplicateurs instables (les multiplicateurs situés à l'extérieur du cercle unité).

La situation hyperbolique (cf. § 14) se présente au voisinage de la trajectoire fermée: toutes les autres trajectoires s'éloignent de la trajectoire fermée aussi bien pour $t \to +\infty$ (l'éloignement a lieu le long de la variété répulsive) que pour $t \to -\infty$ (le long de la variété attractive). Le type topologique de la famille des trajectoires au

voisinage d'une trajectoire fermée, ne possédant pas de multiplicateurs sur le cercle unité, est défini de façon unique par le nombre de multiplicateurs stables et instables et par la parité des nombres de multiplicateurs négatifs contenus parmi les multiplicateurs stables et les multiplicateurs instables.

Voyons quel changement entraîne une petite déformation du système.

B. Singularités les plus simples.

Une trajectoire fermée est *non dégénérée* si l'unité n'est pas multiplicateur. Une trajectoire fermée non dégénérée ne disparaît pas par une petite déformation du système, mais se déforme légèrement (d'après le théorème des fonctions implicites appliqué à l'équation $f(x) = x$, où f est l'application de Poincaré). Les multiplicateurs se déforment légèrement par une déformation de la trajectoire fermée non dégénérée. Donc les nombres de multiplicateurs stables et de multiplicateurs instables ne sont pas modifiés par la déformation si aucun multiplicateur de la trajectoire initiale n'est situé sur le cercle unitaire.

Les multiplicateurs de la trajectoire fermée générique ne sont pas situés sur le cercle unitaire. Donc le système des trajectoires est structurellement stable au voisinage de la trajectoire fermée générique.

Mais si au lieu d'un système nous étudions une famille de systèmes dépendant d'un paramètre, alors pour certaines valeurs du paramètre les multiplicateurs peuvent se retrouver sur le cercle unitaire et poser ainsi le problème des bifurcations.

Comme toujours commençons par les dégénérescences les plus simples, c'est-à-dire par celles qui sont inévitables dans les familles à un paramètre. Ces dégénérescences de codimension un sont ici au nombre de trois. En effet, l'équation caractéristique de l'opérateur de monodromie est réelle, donc chaque multiplicateur non réel possède un conjugué complexe. Par suite, on a sur le cercle unitaire soit deux multiplicateurs conjugués complexes, soit un réel égal à 1 ou à -1. Les trois cas (le couple complexe, $+1$, -1) correspondent à des variétés de codimension 1 dans l'espace fonctionnel.

Considérons, par exemple, la frontière du domaine de stabilité de la trajectoire fermée dans l'espace fonctionnel. Cette frontière est une hypersurface dans l'espace fonctionnel. Elle est constituée de trois composantes de codimension 1. La première composante correspond à une trajectoire ayant un couple de multiplicateurs conjugués complexes de module 1, la deuxième, le multiplicateur $+1$, la troisième, le multiplicateur -1 ; tous les autres multiplicateurs sont situés à l'intérieur du cercle unitaire (fig. 137).

Ces trois hypersurfaces de codimension 1 se coupent suivant des hypersurfaces de codimension 2 et possèdent des singularités ulté-

rieures. Par exemple, les lignes de self-intersection de la première surface correspondent à deux couples de multiplicateurs de module 1, etc.

La perte de stabilité de la trajectoire fermée est donc un problème de dégénérescence des singularités de codimension 1 et il nous faut à première vue considérer des familles génériques à un paramètre pour expliquer les bifurcations. En réalité, la situation n'est pas aussi simple : nous verrons que le problème de la perte de la stabilité au passage d'un couple de multiplicateurs par le cercle unité comporte *deux* paramètres essentiels. Mais voyons tout d'abord à quelles conclusions nous conduit le point de vue « à un paramètre ».

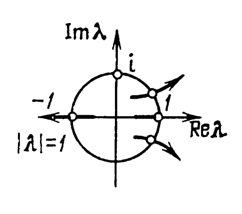


Fig. 137

Commençons par le cas où l'un des multiplicateurs est égal à 1. Ce cas ne diffère pas en principe de la bifurcation des positions d'équilibre dans les familles à un paramètre. La situation générique est la naissance ou la mort d'un couple de trajectoires fermées. Dans ce cas naissent ou meurent deux points fixes de l'application de Poincaré.

Exemple 1. Soit l'application de l'axe x dans lui-même: $x \mapsto x + x^2$. Le point $x = 0$ est fixe et son multiplicateur est égal à un. Considérons la déformation à un paramètre de paramètre ε voisin de 0:

$$f_\varepsilon(x) = x + x^2 + \varepsilon.$$

Cette déformation est topologiquement verselle. Considérons une application quelconque de la droite dans elle-même, possédant un point fixe de multiplicateur 1. On dira que ce point fixe (dégénéré) est *régulier* si la dérivée seconde de l'application en ce point fixe (dans un système de coordonnées quelconque et, par conséquent, dans tout système) est différente de 0.

Si le point fixe non dégénéré est régulier, l'application possède une déformation à un paramètre topologiquement verselle. Ceci étant l'application elle-même et sa déformation verselle sont topologiquement équivalentes localement à l'application spéciale f_0 et à sa déformation f_ε considérée plus haut au voisinage du point 0.

Pour passer au cas multidimensionnel il faut définir une suspension au-dessus de la déformation construite.

Exemple 2. Soit l'application d'un espace vectoriel dans lui-même

$$(y,\ z,\ u,\ v) \mapsto (2y,\ -2z,\ u/2,\ -v/2),$$

où y, z, u, v sont des points des 4 sous-espaces dont le produit direct est l'espace considéré. Nous appellerons une telle application *col standard* (y et u appartiennent à des espaces de dimensions quelconques, z et v, à des espaces de dimensions zéro ou un).

Considérons une application différentiable quelconque possédant un point fixe. Supposons qu'aucun des multiplicateurs n'est situé sur le cercle unitaire. Alors au voisinage de son point fixe l'application est topologiquement équivalente à un col standard (ceci découle aisément du théorème de Grobman-Hartman, § 13).

Exemple 3. Considérons le produit direct d'une déformation de l'application de la droite de l'exemple 1 par un col standard. Nous obtenons une famille à un paramètre d'applications de paramètre ε et de coordonnées variables au voisinage de 0:

$$(x;\ y,\ z,\ u,\ v) \mapsto (x + x^2 + \varepsilon,\ 2y,\ -2z,\ u/2,\ -v/2).$$

Cette déformation est appelée *suspension* au-dessus de la déformation de l'exemple 1. Elle est topologiquement verselle.

Théorème. *Les familles génériques à un paramètre d'applications sont topologiquement équivalentes à la famille mentionnée plus haut au voisinage de chaque point fixe de multiplicateur* 1 *pour les valeurs du paramètre proches de celle pour laquelle le multiplicateur devient égal à* 1.

◀ La démonstration est simple dans le cas unidimensionnel. Le cas multidimensionnel se ramène au cas unidimensionnel à l'aide du théorème de Chochitaïchvili (§ 32) qui est valable non seulement pour les équations différentielles mais aussi pour les applications. ▶

C. Cas du multiplicateur —1.

Dans le cas du multiplicateur —1 la trajectoire fermée dépend différentiablement du paramètre et ne bifurque pas. Mais il en part une trajectoire fermée qui s'enroule deux fois autour d'elle. Pour comprendre ce qui se passe considérons de nouveau l'application de Poincaré.

Exemple 1. Soit l'application d'une droite dans elle-même

$$f_0(x) = -x \pm x^3.$$

Le multiplicateur du point fixe 0 est égal à —1.

Incluons f_0 dans la famille:

$$f_\varepsilon(x) = (\varepsilon - 1)\,x \pm x^3.$$

Théorème. *La déformation f_ε de l'application f_0 est verselle. La famille générique à un paramètre est topologiquement équivalente à la famille mentionnée au voisinage d'un point fixe de multiplicateur —1 pour les valeurs du paramètre proches de celle pour laquelle le multiplicateur est égal à —1.*

◀ Considérons une famille à un paramètre d'applications de la droite dans laquelle le multiplicateur du point fixe est égal à —1 pour une certaine valeur du paramètre.

Le point fixe dépend différentiablement du paramètre (d'après le théorème des fonctions implicites). Par un changement de coordon-

nées dépendant différentiablement du paramètre on peut déplacer le point fixe en 0.

Nous allons maintenant effectuer des changements de Poincaré (cf. § 25) éliminant successivement les termes non résonnants. Ces changements dépendront différentiablement du paramètre si nous conservons les termes qui restent résonnants pour la valeur critique du paramètre non seulement pour cette valeur du paramètre (à ce moment il est impossible de les éliminer), mais aussi pour les valeurs voisines.

Ici les termes résonnants sont tous les termes de puissance impaire. Donc la famille peut être ramenée à la forme

$$x \mapsto \lambda x + ax^3 + O(|x|^5),$$

où λ, a et O dépendent différentiablement du paramètre.

Dans la famille générique la dérivée de λ par rapport au paramètre est différente de 0 pour $\lambda = -1$. Dans ce cas pour paramètre on peut prendre $\varepsilon = 1 + \lambda$. La déformation s'écrit alors

$$x \mapsto (\varepsilon - 1)\, x + a(\varepsilon)\, x^3 + O(|x|^5).$$

Dans la famille générique $a(0) \neq 0$. On peut faire $a(\varepsilon) = \pm 1$ par une dilatation des coordonnées dépendant différentiablement du paramètre.

Il reste maintenant à vérifier que le terme O n'influe pas sur les types topologiques de la famille. Considérons le carré de notre application :

$$x \mapsto (\varepsilon - 1)^2 x + (\varepsilon - 1)\, ax^3 + a(\varepsilon - 1)^3 x^3 + O(|x|^5).$$

Chaque point x se déplace de

$$h = -2\varepsilon(1 + \ldots)\, x - (2a + \ldots)\, x^3 + O(|x|^5),$$

où les points de suspension représentent $O(\varepsilon)$.

Il est facile d'étudier la ligne de niveau nulle de la fonction h sur le plan (x, ε) (fig. 138). La figure 138 définit le type topologique de la famille. ▶

Donc, dans une famille générique à un paramètre des applications d'une droite sur une droite le multiplicateur du point fixe devient égal à -1 au moment du passage transversal par le cercle unité (contrairement au cas où le multiplicateur prend la valeur $+1$ pour laquelle il ne traverse pas généralement le cercle). Lorsque le multiplicateur traverse la valeur -1 de l'intérieur vers l'extérieur, le point fixe perd sa stabilité. Deux cas sont possibles selon le signe du coefficient en x^3. Ou bien on obtient un cycle stable de période 2 (deux points fixes du carré de l'application) à côté du point qui a perdu sa stabilité (à une distance de l'ordre de la racine carrée de l'écart entre le paramètre et la valeur critique) : ce cas correspond à une perte douce de stabilité. Ou bien le rayon du domaine d'attrac-

tion tend vers 0 en raison de l'approche du cycle d'ordre 2 avant la perte de stabilité (perte brusque de stabilité).

Le cas multidimensionnel se déduit par suspension d'un col comme précédemment.

En transposant tout ce qui a été dit sur les applications à l'application de Poincaré de la trajectoire fermée, on obtient dans le cas

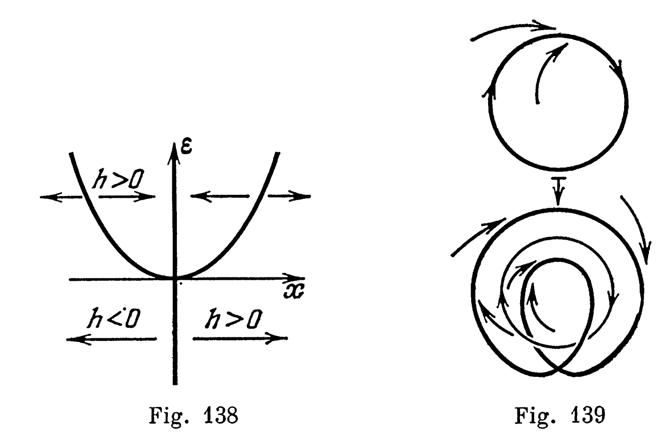


Fig. 138 Fig. 139

d'une perte douce de stabilité le tableau représenté sur la figure 139 : le cycle initial perd sa stabilité mais il apparaît un cycle stable avec une période approximativement deux fois plus longue.

Les phénomènes décrits ici sont bien observables dans les expériences. L'exemple qui va suivre a été emprunté à une conférence de G. Barenblat au séminaire de I. Pétrovski. On considère une pellicule polymère distendue lentement par une charge. Le processus est quasi stationnaire pour de petites tensions (on peut admettre que le temps est un paramètre, que le point représentatif se trouve dans une position d'équilibre stable et que toutes les grandeurs observées pour chaque valeur du paramètre sont constantes, c'est-à-dire varient lentement avec le temps). Cependant pour une certaine valeur du paramètre (c'est-à-dire pour une distension suffisante de la pellicule) le tableau change et les divers paramètres physiques (disons la longueur x de la pellicule) traités comme fonction du temps prennent la forme représentée sur la figure 140 (chaque oscillation peut être considérée comme se produisant pour une valeur fixe du paramètre, mais que ce paramètre est un peu différent pour l'oscillation suivante).

Voici l'interprétation de ce comportement des variables avec le temps: le point 1 correspond à une perte douce de stabilité de la position d'équilibre avec génération d'auto-oscillation; on voit que leur amplitude croît comme la racine carrée de l'écart entre le paramètre et la valeur critique pour laquelle a lieu la bifurcation. Le point 2 correspond à une perte douce de stabilité du cycle avec passage du multiplicateur par la valeur —1.

En effet, supposons que les modifications indiquées sur la figure 139 ont lieu dans l'espace des phases.

Chaque grandeur physique observable est une fonction sur l'espace des phases. Cette grandeur est constante tant que le point représentatif se trouve dans la position d'équilibre. Lorsque le point représentatif se déplace le long du cycle, la grandeur x devient une fonction périodique du temps t (l'amplitude des oscil-

lations croît avec le cycle). Au doublement du cycle, représenté sur la figure 139, correspond précisément un doublement de la période de la dépendance de la grandeur mesurée par rapport au temps, observé au cours de l'expérience (fig. 140).

On remarquera de ce fait que lorsqu'on étudie des auto-oscillations, on enregistre d'habitude les dépendances des grandeurs mesurées par rapport au temps (par exemple sur un électrocardiogramme). Dans de nombreux cas l'allure

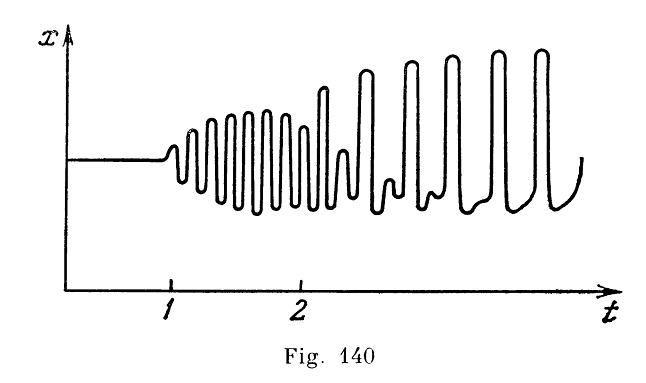


Fig. 140

de la trajectoire dans l'espace des phases ou de sa projection sur un plan quelconque nous éclaire mieux sur le caractère des phénomènes. Ces méthodes sont utilisées depuis longtemps pour déceler les pannes de systèmes mécaniques auto-oscillants tels que les pompes. Les médecins ont proposé d'appliquer cette méthode en électrocardiographie.

D. Passage d'un couple de multiplicateurs par le cercle unité. Ce cas est bien moins étudié que les deux précédents. Les déformations topologiquement verselles n'ont pas été explicitées et il est même possible qu'elles n'existent pas. Cependant la méthode de Poincaré permet de recueillir une information importante. Commençons par le cas où l'argument du multiplicateur situé sur le cercle unité n'est pas commensurable avec 2π (on peut admettre que ce cas est générique, car l'ensemble des nombres rationnels est de mesure nulle).

On admettra que l'espace appliqué est de dimension 2. Notre famille d'applications se ramène alors par un changement convenable de coordonnées différentiables et dépendant différentiablement du paramètre à la forme

$$z \mapsto \lambda(\varepsilon) z (1 + a(\varepsilon) |z|^2 + O(|z|^4)),$$

où le nombre réel ε est le paramètre de la famille, $\lambda(0) = e^{i\alpha}$, $\alpha \neq 2\pi p/q$. Pour la famille générique on a $d|\lambda|/d\varepsilon|_0 \neq 0$, de sorte que l'on peut prendre $|\lambda| - 1$ pour paramètre.

Supposons que le terme $O(|z|^4)$ ne figure pas dans la formule. L'application est alors facile à étudier. En effet, le module de l'image d'un point est défini par celui de l'antécédent, de sorte que l'on

obtient l'application réelle

$$r \mapsto r\,|\,\lambda\,|\,|\,1 + ar^2\,|.$$

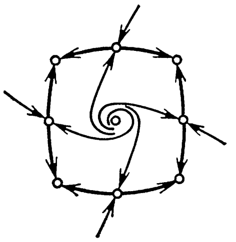
Pour $|\,\lambda\,| = 1 + \varepsilon$, $|\,\varepsilon\,| \ll 1$, $r \ll 1$, on a

$$|\,\lambda\,|\,|\,1 + ar^2\,| \approx 1 + \varepsilon + \operatorname{Re} ar^2 + \ldots$$

Pour la famille générique on a $\operatorname{Re} a \neq 0$. Dans ce cas lorsque le paramètre ε prend la valeur 0, du point fixe perdant sa stabilité il naît (ou bien en ce point il meurt) un cercle invariant par cette application, de rayon proportionnel à $\sqrt{|\,\varepsilon\,|}$. Dans le premier cas (la naissance) ce cercle est stable, dans le second, il est instable. Sur le cercle l'application se réduit à une rotation.

Revenons maintenant aux termes négligés et voyons s'ils modifient nos conclusions.

On démontre que l'application complète possède effectivement une courbe fermée invariante de rayon d'ordre $\sqrt{|\,\varepsilon\,|}$ (voir Sacker R. J., On invariant surfaces and bifurcation of periodic solutions of ordinary differential equations, New York University, Report IMM — NYU 333, 1964; CPAM 18, 4 (1965), 717-732).


Fig. 141

Cette courbe fermée reste également stable par une perturbation. Cependant la structure de l'application complète sur cette courbe invariante diffère de celle de l'application privée de son résidu. En effet, sur la courbe invariante l'application complète peut posséder un nombre de rotation aussi bien irrationnel que rationnel. L'application du cercle obtenue n'est pas forcée d'être topologiquement équivalente à une rotation. Dans le cas du nombre rationnel de rotation, elle aura en général un nombre fini de points périodiques alternativement stables et instables. Ces points périodiques seront respectivement des cols et des nœuds pour l'application initiale du plan lui-même. Donc, si le nombre de rotation est rationnel, notre courbe invariante sera constituée d'une chaîne de séparatrices de cols convergeant dans les nœuds (fig. 141).

A noter que les séparatrices des cols sont lisses. Cependant à l'abord d'un nœud des deux côtés deux séparatrices constituent ensemble, généralement, une courbe différentiable un nombre fini de fois. Par suite, la courbe invariante engendrée n'est différentiable qu'un nombre fini de fois. Lorsque le paramètre tend vers une valeur correspondant au passage des multiplicateurs par le cercle, la courbe invariante tend à devenir indéfiniment différentiable (ce qui est aisé à voir).

Si notre application est une application de Poincaré pour une équation différentielle, la courbe invariante de l'application de

Poincaré définit dans l'espace des phases à trois dimensions un tore invariant sur lequel les trajectoires sont partout denses. Notre courbe invariante est la section transversale de ce tore. Le tore appartient à la classe C^r (r fini), et r est d'autant plus grand que l'instant de la naissance du tore à partir du cycle est proche. Une variation du paramètre de la famille entraîne en général une variation du nombre de rotation sur le tore, de sorte que ce nombre prend une valeur tantôt irrationnelle, tantôt rationnelle.

L'étude des bifurcations correspondant au passage d'un couple de multiplicateurs par le cercle unitaire nous indique aussi que les familles génériques à un paramètre ne donnent pas lieu à des ramifications du régime périodique donné en régimes périodiques de multiplicité distincte de deux. En effet, ce cas se serait présenté seulement si le multiplicateur avait traversé le cercle unité en un point d'argument rationnel, or ce phénomène est exceptionnel.

Pour comprendre comment apparaissent les mouvements périodiques de périodes élevées, il est indispensable d'étudier les familles à deux paramètres.

En effet, le cas où le multiplicateur est égal à la racine de l'unité (distinct de 1 et —1) est inévitable seulement dans les familles réelles à deux paramètres. L'étude « à deux paramètres » de la perte de stabilité du point fixe dans le cas résonnant, c'est-à-dire lorsque le multiplicateur est proche d'une racine de l'unité, permet également de mieux comprendre les bifurcations dans les familles à un paramètre lorsque les multiplicateurs traversent le cercle unité. Nous verrons notamment que certaines modifications semblant non locales dans l'approche « à un paramètre » se laissent étudier par des méthodes locales si le problème est traité comme un problème à deux paramètres. En particulier, dans cette optique on peut étudier certains cas de perte brusque de stabilité et indiquer quel régime sera adopté ensuite par le système.

E. Résonance pendant une perte de stabilité du cycle.

Considérons une application du plan dans lui-même au voisinage d'un point fixe de multiplicateur égal à la racine q-ième ($q > 2$) de l'unité. Suivant la méthode générale de Poincaré (chap. 5) nous pouvons écrire la famille dans un système de coordonnées convenables sous la forme

$$z \mapsto z[\lambda + A(|z|^2) + B\bar{z}^{q-1} + O(|z|^{q+1})],$$

où λ, A, B et O dépendent différentiablement de ε.

Au lieu d'étudier cette application nous allons procéder autrement. En cas de résonance chaque pas de la méthode de Poincaré se ramène à une moyennisation le long du feuilletage correspondant de Seifert (cf. § 21). Donc au lieu de réduire l'application de Poincaré à la forme normale, on peut écrire l'équation initiale des trajectoires

du cycle comme une équation non autonome à coefficients 2π-périodiques et ensuite la réduire à la forme normale par des changements de coordonnées $2\pi q$-périodiques en t (cf. § 26).

On obtient par cette procédure une équation à coefficients $2\pi q$-périodiques en t dans les nouvelles coordonnées (dépendant différentiablement du paramètre):

$$\dot{\zeta} = \varepsilon\zeta + \zeta A(|\zeta|^2) + B\bar{\zeta}^{q-1} + O(|\zeta|^{q+1}),$$

où ε est un paramètre complexe, A et B dépendent holomorphement de ε, la valeur $\varepsilon = 0$ correspond à une résonance (c'est-à-dire à l'égalité du multiplicateur de l'équation initiale à la racine q-ième de 1).

Remarque 1. Ces raisonnements entraînent, en particulier, que:

1) L'application de Poincaré est confondue aux termes d'ordre $q + 1$ près (et même aux termes d'ordre aussi grand que l'on veut près) avec une application du flot du champ de vecteurs sur le plan.

2) Le champ de vecteurs indiqué est invariant par le groupe cyclique des difféomorphismes du plan (d'ordre q).

3) Les conclusions 1 et 3 sont valables non seulement pour notre application de Poincaré mais aussi pour une famille dépendant différentiablement de paramètres, et de plus le groupe et le champ invariant par groupe dépendent différentiablement des paramètres.

Remarque 2. L'application exacte de Poincaré n'est en général application du flot d'aucun champ de vecteurs et ne commute avec aucun groupe fini de difféomorphismes.

De ce qui précède on voit que le problème des bifurcations lors d'une perte de stabilité au voisinage d'une résonance d'ordre $q > 2$ se ramène, aux termes d'ordre aussi grand que l'on veut près par rapport à la distance à la trajectoire fermée, à l'étude des modifications des portraits de phases dans des familles génériques à deux paramètres de champs de vecteurs sur le plan, invariants par des rotations d'angle $2\pi/q$. La résonance est dite *forte* si $q \leqslant 4$.

Le cas des résonances d'ordres 2 et 1 peut être inclu dans ce schéma. Plus exactement, la perte de stabilité du cycle au passage d'un couple de multiplicateurs par le cercle unitaire correspond à une hypersurface de codimension 1 dans l'espace fonctionnel. Cette hypersurface aborde les hypersurfaces correspondant aux multiplicateurs 1 et —1 suivant deux surfaces de codimension 2. Les points génériques de ces surfaces de codimension 2 correspondent à des trajectoires fermées telles que l'application de Poincaré possède une valeur propre double 1 (respectivement —1) de bloc de Jordan d'ordre 2.

Donc, l'étude des cas frontières de passage des multiplicateurs par le cercle unitaire se ramène aux termes d'ordre aussi grand que l'on veut près, à celle des modifications des portraits de phases dans les familles génériques à deux paramètres de champs de vecteurs sur le plan, invariants par les rotations d'angle $2\pi q$ ($q = 1, 2$) et possé-

dant pour une certaine valeur du paramètre un point singulier dont la partie linéaire est de la forme d'un bloc de Jordan nilpotent; l'équation linéaire correspondante se ramène à la forme

$$\dot{x} = y, \qquad \dot{y} = 0.$$

Finalement, le problème des modifications pendant une perte de stabilité au voisinage des résonances se ramène à l'étude des bifurcations des portraits de phases dans les familles à deux paramètres de champs de vecteurs équivariants sur le plan; c'est ce problème que nous allons étudier maintenant.

§ 35. Déformations verselles de champs de vecteurs équivariants sur le plan

Les modifications des portraits de phases des champs de vecteurs invariants par un groupe quelconque de symétries se présentent naturellement quand on étudie divers phénomènes dans lesquels la symétrie figure dans la position même du problème.

Le plus étonnant est que les problèmes de modifications des portraits de phases symétriques se posent dans une situation à priori non symétrique lorsqu'on étudie les bifurcations au voisinage des résonances (cf. §§ 21 et 34). Dans ce paragraphe on considère notamment les bifurcations des portraits de phases symétriques indispensables à l'étude des résonances.

A. Champs de vecteurs équivariants sur le plan.

Soit F un champ de vecteurs sur le plan de la variable complexe z. On traitera F comme une fonction à valeurs complexes (pas forcément holomorphe) sur $\mathbb{C}$. La série de Taylor de cette fonction s'écrit en 0: $\sum F_{k,l}\, z^k \bar{z}^l$.

P r o p o s i t i o n. *Supposons que le champ F se transforme dans lui-même par une rotation d'angle $2\pi/q$ du plan de la variable z. Alors les coefficients de $F_{k,l}$ ne sont pas nuls uniquement pour les $k - l$ congrus 1 modulo q.*

◀ La série de Taylor est unique, donc chacun de ses termes doit pivoter d'un angle de $2\pi/q$ lorsque z pivote de $2\pi/q$. Le point du plan complexe $z^k\bar{z}^l$ pivote d'un angle de $2\pi\,(k - l)/q$. Cette rotation coïncide avec une rotation d'angle $2\pi/q$ sous la condition indiquée plus haut. ▶

C o r o l l a i r e. *Les équations différentielles invariantes par les rotations d'angle $2\pi/q$ sont de la forme:*

$$\dot{z} = zA(|z|^2) + B\bar{z}^{q-1} + O(|z|^{q+1}) \qquad (q > 2).$$

◀ Considérons sur le plan $(k,\ l)$ les points entiers tels que $k - l \equiv 1 \pmod{q}$. Ces points sont situés sur des rayons parallèles à la

bissectrice du premier quadrant, issus des points correspondant aux monômes z, z^{q+1}, z^{2q+1}, . . .; $\bar{z}^{q-1}$, $\bar{z}^{2q-1}$, . . . Parmi ces monômes on cherchera ceux qui sont de plus petits degrés (fig. 142).

On obtient successivement d'abord quelques monômes sur le rayon issu de z (c'est-à-dire des monômes de la forme $z\,|\,z\,|^{2q}$), ensuite le monôme $\bar{z}^{q-1}$; tous les autres nonômes sont de degré $\geqslant q+1$ (fig. 142). ▶

D é f i n i t i o n. On appelle *équation principale* invariante par une rotation d'angle $2\pi/q$ l'équation précédente privée du terme O.

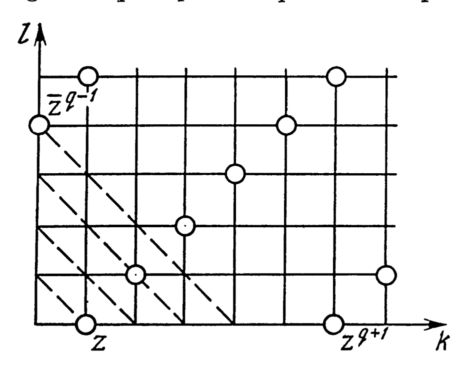


Fig. 142

Le second membre de l'équation principale est appelé *champ q-équivariant principal*.

E x e m p l e. Les équations principales invariantes par les groupes de rotation d'ordre 3 et 4 s'écrivent respectivement

$$\dot{z}=\varepsilon z+Az\,|z|^2+Bz^2,\quad \dot{z}=\varepsilon z+Az\,|z|^2+B\bar{z}^3.$$

Il est commode d'introduire les définitions suivantes pour formuler les résultats de l'étude des modifications des portraits de phases de champs de vecteurs équivariants, dépendant de paramètres.

B. Déformations verselles équivariantes.

Considérons une famille v_λ de champs de vecteurs invariants par un groupe G dans l'espace des phases et dépendant d'un paramètre λ appartenant au voisinage du point 0 de l'espace $\mathbb{R}^k$ (appelé *base* de la famille). La dimension de la base est appelée *nombre de paramètres* de la famille.

On appelle *déformation équivariante* du champ v_0 le germe de la famille au point $\lambda=0$.

D é f i n i t i o n. On dit qu'une déformation équivariante v_λ est une *déformation verselle* v_0 si pour toute autre déformation équivariante w_μ du même champ il existe une application continue φ

des bases des déformations et une famille d'homéomorphismes h_μ de l'espace des phases, dépendant continûment de μ commutant avec l'action de G, telle que h_μ envoie les trajectoires du champ w_μ dans les trajectoires du champ $v_{\varphi(\mu)}$ en conservant le sens du mouvement.

Il est équivalent de dire *qu'une déformation équivariante est verselle si toute autre déformation équivariante est topologiquement orbitalement équivalente à une déformation induite par la déformation verselle.*

On introduit des définitions analogues pour les germes de champs de vecteurs ainsi que pour les déformations dans une classe de champs possédant des propriétés spéciales (par exemple à partie linéaire fixe en un point singulier). Passons maintenant à la construction des déformations verselles.

C. Déformations principales.

Considérons un champ de vecteurs sur le plan, invariant par une rotation d'angle $2\pi/q$, $q > 2$.

D é f i n i t i o n. Ce champ est *singulier* si sa partie linéaire est nulle au point 0.

D é f i n i t i o n. On appelle *déformation principale d'un champ singulier principal q-équivariant* v_0 ($q > 2$) la famille à deux paramètres $v_\varepsilon = \varepsilon z + v_0$, où les paramètres sont les parties réelle et imaginaire du nombre complexe ε.

E x e m p l e. Lorsque $q = 3$ et 4, les déformations principales sont données par les équations

$$\dot{z} = \varepsilon z + Az|z|^2 + B\bar{z}^2, \quad \dot{z} = \varepsilon z + Az|z|^2 + B\bar{z}^3,$$

où ε est traité comme un paramètre, et les coefficients complexes A et B sont fixes.

R e m a r q u e. On a affaire aux objets définis plus haut lorsqu'on étudie la perte de stabilité d'un cycle au passage d'un couple de multiplicateurs conjugués complexes par le cercle unitaire. Les systèmes comportant un tel passage forment une hypersurface dans l'espace fonctionnel de tous les systèmes. Cette hypersurface est l'une des trois hypersurfaces limitant le domaine de stabilité. Les deux autres hypersurfaces correspondent au passage d'un multiplicateur par le cercle unitaire au point 1 et au point —1 respectivement.

La frontière de l'hypersurface correspondant au passage d'un couple complexe de multiplicateurs est constituée de deux parties (deux surfaces de codimension 2 dans l'espace fonctionnel de tous les systèmes). L'une de ces surfaces de codimension 2 correspond à un couple de multiplicateurs égaux à 1 et formant un bloc de Jordan d'ordre 2, l'autre, à un même bloc possédant deux valeurs propres égales à —1.

L'étude de la perte de stabilité au voisinage de ces surfaces limites de codimension 2 revient à celle des bifurcations dans des familles à deux paramètres de champs de vecteurs sur le plan, dont la partie linéaire est un bloc de Jordan nilpotent d'ordre 2 et qui sont symétriques par les rotations d'angle $2\pi/q$, $q = 2$ (pour les multiplicateurs -1) ou $q=1$ (pour les multiplicateurs 1). Pour inclure l'étude de ces cas dans le schéma général on aura besoin des définitions suivantes.

D. Cas $q = 1$ et $q = 2$.

D é f i n i t i o n. Un champ invariant par une rotation du plan d'angle $2\pi/q$, $q = 1$ ou 2, est *singulier* si sa partie linéaire est un bloc de Jordan nilpotent d'ordre 2 au point 0.

En d'autres termes, pour $q = 1$ ou 2, le champ singulier est un champ dont la partie linéaire est le champ de la vitesse de l'équation $\ddot{x} = 0$ sur le plan de phase $(x, y = \dot{x})$.

On démontre aisément le

T h é o r è m e. *Un champ singulier invariant par une rotation du plan d'angle $2\pi/q$, $q = 1$ ou 2, se ramène par un difféomorphisme commutant avec cette rotation au champ de la vitesse de l'équation*

$$\ddot{x} = ax^3 + bx^2y + O(|x|^5, |y|^5) \quad (q=2),$$
$$\ddot{x} = ax^2 + bxy + O(|x|^3, |y|^3) \quad (q=1)$$

sur le plan de phase $(x, y = \dot{x})$.

◀ La partie linéaire de notre champ est de la forme $y\partial/\partial x$. Composons l'équation homologique correspondant à ce champ linéaire. Calculons à cet effet le crochet de Poisson de notre champ linéaire $\Lambda = y\partial/\partial x$ et d'un champ de vecteurs arbitraires $h = P\partial/\partial x + Q\partial/\partial y$.

On a successivement

$$[y\partial/\partial x, P\partial/\partial x] = yP_x\partial/\partial x, \quad [y\partial/\partial x, Q\partial/\partial y] = yQ_x\partial/\partial y - Q\partial/\partial x,$$
$$[\Lambda, h] = (yP_x - Q)\,\partial/\partial x + yQ_x\partial/\partial y.$$

Donc, l'équation homologique en les fonctions inconnues (P, Q) a la forme du système

$$yP_x - Q + u = 0, \quad yQ_x + v = 0,$$

où u et v sont des fonctions inconnues, plus exactement les composantes du champ de vecteurs $w = u\partial/\partial x + v\partial/\partial y$ que nous voulons éliminer par un changement de variables.

Etudions le système obtenu. En éliminant Q de la première équation et en le portant dans la seconde, on obtient

$$y^2P_{xx} = -yu_x - v.$$

Pour que le second membre soit divisible par y^2 il suffit de changer les termes de degrés 0 et 1 en y de la fonction v. Si donc l'on

substitue $v_0(x) + yv_1(x)$ à v, on peut résoudre la dernière équation par rapport à P.

Donc l'équation homologique insoluble pour (u, v) quelconques devient soluble si l'on remplace v par une fonction linéaire non homogène convenable en y. Autrement dit, l'équation

$$[\Lambda, h] + w = (v_0(x) + yv_1(x))\,\partial/\partial y$$

est soluble par rapport au champ inconnu h pour (v_0, v_1) convenables, dépendant de w.

En définitive, la méthode de Poincaré permet d'éliminer dans les termes de chaque puissance du champ de vecteurs $\Lambda + \ldots$ tous les monômes vectoriels à l'exception de ceux de la forme $x^k\partial/\partial y$ et $yx^k\partial/\partial y$. Donc, dans la classe des séries entières formelles notre équation se ramène à la forme

$$\ddot{x} = a(x) + yb(x).$$

Si le système initial était invariant par une rotation d'angle π (c'est-à-dire était impair), c'est que les composantes du champ de vecteurs initial étaient des fonctions impaires. Dans ce cas les changements de la méthode de Poincaré peuvent être choisis également impairs (commutant avec la rotation), puisque les puissances de (P, Q) et (u, v) sont égales dans les formules précédentes. Les séries de (a, b) dans la forme normale formelle seront alors composées uniquement de termes d'ordre impair.

En se limitant dans la méthode de Poincaré à quelques premières approximations on obtient la proposition formulée plus haut.

D é f i n i t i o n. On appelle *équations et champs principaux singuliers* pour $q = 2$ et 1 respectivement les équations

$$\ddot{x} = ax^3 + bx^2y \quad (q = 2), \qquad \ddot{x} = ax^2 + bxy \quad (q = 1)$$

et les champs de vecteurs qui les définissent sur le plan de phase $(x, y = \dot{x})$.

D é f i n i t i o n. On appelle *déformation principale* d'un champ singulier principal pour $q = 2$ et 1 respectivement la déformation consistant à ajouter au second membre de l'équation de second ordre le terme $\alpha x + \beta y$ $(q = 2)$ et $\alpha + \beta x$ $(q = 1)$.

Voici la liste des déformations principales de champs q-équivariants dans les cas d'une résonance forte, c'est-à-dire pour $q \leqslant 4$:

$$\dot{z} = \varepsilon z + Az|z|^2 + B\bar{z}^3, \qquad q = 4,$$

$$\dot{z} = \varepsilon z + Az|z|^2 + B\bar{z}^2, \qquad q = 3,$$

$$\ddot{x} = \alpha x + \beta y + ax^3 + bx^2y, \qquad q = 2,$$

$$\ddot{x} = \alpha + \beta x + ax^2 + bxy, \qquad q = 1,$$

où les variables z, ε, A, B sont complexes; x, y, α, β, a, b, réelles; les paramètres de la déformation sont désignés par des lettres grecques; $y = \dot{x}$.

E. Versalité des déformations principales.

« Théorème ». *Pour chaque q tous les champs singuliers principaux peuvent être décomposés en dégénérés et non dégénérés de telle sorte que*

1. *Les champs dégénérés forment une réunion d'un nombre fini de sous-variétés dans l'espace des champs singuliers principaux.*

2. *Les champs non dégénérés forment une réunion d'un nombre fini de domaines connexes ouverts.*

3. *La déformation principale du germe d'un champ non dégénéré est verselle en* 0.

4. *Les déformations principales des germes des champs non dégénérés sont topologiquement équivalentes dans chaque composant de connexité.*

On a mis le mot « théorème » entre guillemets, car il n'a pas été démontré pour $q = 4$.

A l'exception du cas $q = 4$, les conditions de non-dégénérescence se détaillent comme suit:

$$a \neq 0,\ b \neq 0 \text{ pour } q = 1,\ 2;\ \operatorname{Re} A\,(0) \neq 0,\ B \neq 0 \text{ pour } q \geqslant 3.$$

Pour $q = 4$ il faut ajouter à ces conditions au moins les conditions (toutes les autres sont probablement indiquées aux p. p. 318-319, 3), 4)):

$$|A|^2 \neq |B|^2, \quad |\operatorname{Re} A| \neq |B|,$$
$$|\operatorname{Im} A| \neq (|B|^2 + \operatorname{Re}^2 A)/\sqrt{|B|^2 - \operatorname{Re}^2 A}.$$

Ce théorème entraîne immédiatement le

Corollaire. *Dans l'espace fonctionnel des familles à deux paramètres de champs de vecteurs, invariants par le groupe de rotations d'angles multiples de $2\pi/q$, les familles pour lesquelles les champs sont singuliers uniquement pour certaines valeurs isolées des paramètres et au voisinage de ces valeurs sont topologiquement équivalentes à des déformations verselles des champs singuliers non dégénérés principaux, forment un ensemble partout dense ouvert* *).

En d'autres termes, supposons que les valeurs propres du linéarisé d'un champ de vecteurs sur le plan, invariant par les rotations d'angle $2\pi/q$, sont nulles. Considérons le champ générique jouissant de ces propriétés. Composons sa déformation à deux paramètres générique dans la classe des champs possédant la même symétrie.

On affirme qu'il existe un voisinage du point 0 ne dépendant pas des paramètres tel que la déformation construite se réduise à la forme normale indiquée dans les points C et D par un homéomorphisme dépendant continûment des paramètres et possédant la

*) Avec les restrictions habituelles si la base n'est pas compacte.

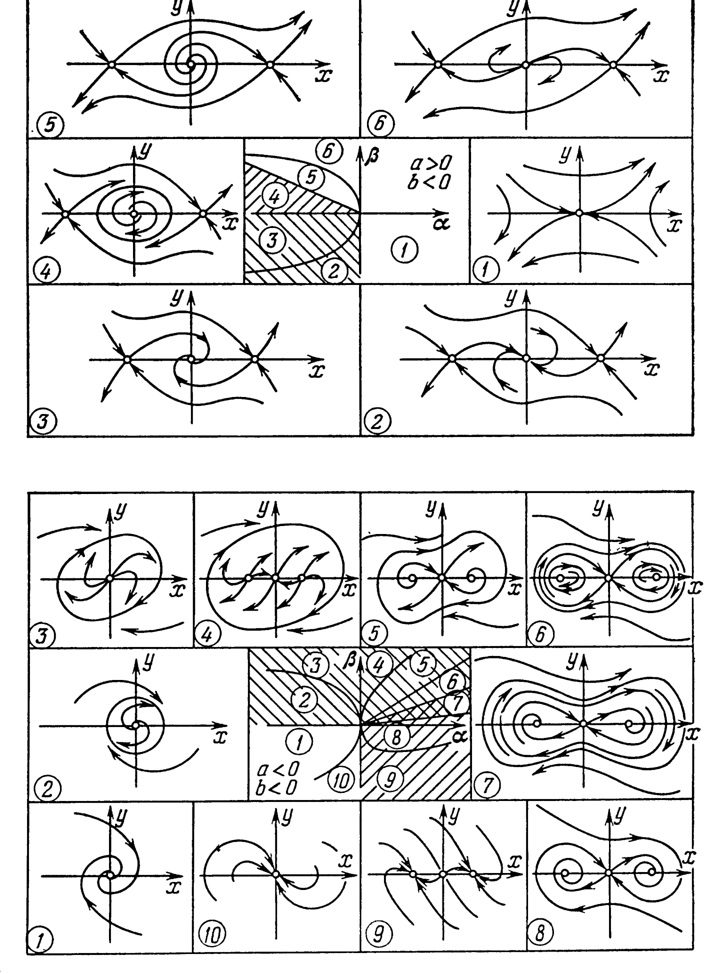


Fig. 143

même symétrie. Plus exactement, les homéomorphismes réduisent les portraits de phases des systèmes correspondants à la forme normale.

Donc, le théorème formulé ramène la description de toutes les bifurcations à celle des bifurcations dans les déformations principales de champs singuliers non dégénérés.

F. Description des bifurcations.

Pour $q = 1$ le théorème formulé plus haut a été démontré par R. Bogdanov en 1971 (cf. § 33, où l'on trouvera la description de ces bifurcations).

Fig. 144

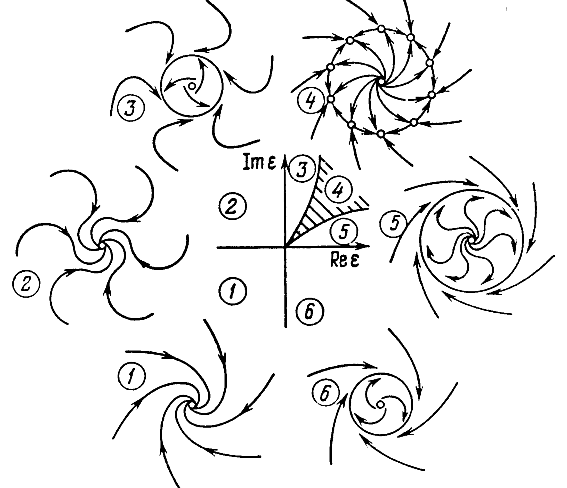


Fig. 145

Pour $q = 2$ on réalise $b < 0$ par des changements du temps. Le diagramme des bifurcations (le « cadran ») sur le plan (α, β) et la modification du portrait de phases pour $a > 0$ et $a < 0$ sont représentés sur la figure 143.

Pour $q = 3$ on rend $\operatorname{Re} A < 0$ par des changements du temps. Le diagramme des bifurcations et la modification sont représentés sur la figure 144.

Pour $q \geqslant 5$ on s'arrange pour que $\operatorname{Re} A < 0$ par des changements du temps; le diagramme des bifurcations et les modifications sont représentés sur la figure 145. On remarquera que la zone d'existence de points fixes se rapproche de l'axe imaginaire ε sous la forme d'une langue étroite dont les extrémités possèdent une tangente commune:

$$\operatorname{Im} \varepsilon \approx f\,|\operatorname{Re} \varepsilon| \pm c\,|\operatorname{Re} \varepsilon|^{(q-2)/2}, \quad q \geqslant 5.$$

Les démonstrations du théorème et des propositions citées plus haut sont simples pour $q \geqslant 5$, sont contenues dans les travaux de Bogdanov cités au § 33 pour $q = 1$, n'ont pas été effectuées pour $q = 4$ et sont ébauchées plus bas pour $q = 2$ et 3. Certaines variantes des modifications pour $q = 4$ sont représentées plus bas (sur les figures 149, 150, 152). Voir également aux p.p. 318-319, 1), 2), 3), 4).

G. Cas d'une symétrie d'ordre 3.

1°. Si $A = 0$, le système se déduit à partir d'un système hamiltonien par une rotation du champ. Plus exactement, considérons le triangle équilatéral formé par les points singuliers (cols). Les côtés de ce triangle définissent trois fonctions non homogènes linéaires. Le produit de ces trois fonctions n'est autre que la fonction de Hamilton du système.

Dans notre système $\dot{z} = \varepsilon z + B\bar{z}^2$, le signe de la dérivée de cette fonction, est défini par celui de la partie réelle du paramètre ε. Cette circonstance nous permet d'utiliser cette fonction comme fonction de Liapounov. Donc, pour $A = 0$ l'étude de la déformation principale se fait sans aucune difficulté.

2°. Dans le cas général les substitutions $t = T/|\varepsilon|$, $z = Z|\varepsilon|$ ramène le système à la forme

$$dZ/dT = EZ + \bar{Z}^2 + |\varepsilon|\, AZ\,|Z|^2, \quad E = \varepsilon/|\varepsilon|.$$

Dans le domaine où $|Z|$ est petit en regard de $1/\varepsilon$ le troisième terme peut être considéré comme une petite perturbation. Dans le domaine où l'argument de E n'est pas proche de $\pm\pi/2$, cette perturbation ne modifie pas le tableau du 1°. Si la partie réelle de E est petite en regard de la partie complexe, on peut considérer le système comme un système hamiltonien peu perturbé $dZ/dT = \pm iZ + \bar{Z}^2$.

Le hamiltonien non perturbé H est décrit au 1°.

3°. Calculons la vitesse de variations de H le long des solutions du système. Une intégration sur les lignes de niveau $H = h$ nous donne la condition de naissance d'un cycle sur cette ligne sous la forme

$$\oint \mu r^2 + \lambda r^4 d\varphi = 0,$$

où $\mu = -\sigma/\tau$, $\lambda = a\tau$, $\varepsilon = \sigma + i\tau$, $A = -a + ib$; r et φ sont les coordonnées polaires sur le plan.

Désignons par ρ le rayon d'inertie du domaine limité par la courbe elliptique $H = h$. La condition de naissance d'un cycle à partir précisément de la ligne $H = h$ au passage de $|\varepsilon|$ par 0 peut alors être écrite sous la forme : $\sigma = \rho^2\tau^2 a$.

La valeur maximale de ρ correspond au triangle formé par les séparatrices des cols. La fonction $\rho(h)$ est monotone (cf. I).

4°. La justification de la versalité de la famille principale et de la forme des diagrammes des bifurcations et des portraits de phases s'achève comme pour $q = 1$.

H. Cas d'une symétrie d'ordre 2.

1°. Par des dilatations de x et des changements du temps et des paramètres, on ramène la famille à la forme

$$\ddot{x} = \alpha x + 2\beta y + a x^3 + b x^2 y, \quad a = \pm 1, \quad b = -2.$$

Nous étudierons cette famille de paramètres (α, β).

2°. Si $|\beta| \leqslant \sqrt{\alpha}$, on effectue le changement $x = \sqrt{|\alpha| / |a|}\, x'$, $t = t' / \sqrt{|\alpha|}$ qui ramène $|\alpha|$ et $|a|$ à 1. On obtient un système presque hamiltonien de fonction de Hamilton

$$H = \frac{y^2}{2} - \operatorname{sgn} \alpha \frac{x^2}{2} - \operatorname{sgn} a \frac{x^4}{4}.$$

Les termes dissipatifs sont de la forme

$$2\beta' y + b' x^2 y, \quad \beta' = \beta / \sqrt{|\alpha|}, \quad b' = \sqrt{|\alpha|}\, b / |a|.$$

3°. En intégrant la vitesse de variation de H le long des lignes de niveau $H = h$, on obtient la condition de naissance d'un cycle à partir précisément de cette ligne au passage de α et $\beta = u\alpha$ par 0 : $u = r^2$, où

$$r^2 = \iint x^2\, dx\, dy \Big/ \iint dx\, dy$$

est le carré du rayon d'inertie par rapport à l'axe y du domaine limité par la ligne $H = h$.

4°. Si $|\alpha| \ll |\beta|$, on fait la substitution $x = \lambda x'$, $t = \varkappa t'$ qui ramène $|a|$ à 1 et $|b|$ à $|\beta|$:

$$\lambda = \sqrt{|\beta / b|}, \quad \varkappa = \sqrt{|b / \beta|}.$$

Il vient $\alpha' = |b/\beta|\, \alpha$, $b' = \sqrt{|\beta / b|}\, b$, $\beta' = \sqrt{|b / \beta|}\, \beta$, $a' = a$;

$$\ddot{x} = a x^3 + \beta' 2y\, (1 \pm x^2) + \alpha' x,$$

les deux paramètres $\beta' \sim \sqrt{|\beta|}$ et $\alpha' \sim \alpha/\beta$ sont petits. Pour $\beta' = 0$ on obtient le système hamiltonien

$$H = \frac{y^2}{2} - \frac{\alpha' x^2}{2} - \frac{a x^4}{4}.$$

L'étude s'achève comme de coutume.

I. Zéros des intégrales elliptiques.

On a vu plus haut que l'étude du comportement des cycles dans nos familles s'est ramenée à la résolution de cas spéciaux du « 16-ième problème faible de Hilbert » suivant :

Soient H un polynôme réel de degré n, P un polynôme réel de degré m en (x, y). *Déterminer le nombre de zéros réels de la fonction*

$$I(h) = \iint_{H \leqslant h} P\, dx\, dy.$$

Pour étudier les symétries d'ordre 2 il faut que

$$P = \mu + \lambda x^2, \quad H = \frac{y^2}{2} - \frac{\alpha x^2}{2} - \frac{a x^4}{4}.$$

Pour $P = \mu + \lambda x^2$ le problème se ramène à l'étude de la monotonie d'une fonction $r(h)$, où r est le rayon d'inertie par rapport à l'axe y du domaine borné par le cycle.

Lemme. *La fonction r se conduit de la manière suivante dans les intervalles de valeurs critiques de H*:

Valeurs de α et a	$-1, +1$	$+1, -1$	$+1, -1$	$-1, -1$
Intervalle de h	$0, \frac{1}{4}$	$-\frac{1}{4}, 0$	$0, +\infty$	$0, +\infty$
Comportement de r	$\uparrow$	$\downarrow$	$\downarrow \uparrow$	$\uparrow$

(*dans le troisième cas, r décroît, ensuite croît*).

R. Bogdanov a utilisé dans ses travaux un lemme analogue (mais plus faible) sur les intégrales elliptiques. La démonstration de R. Bogdanov comporte des calculs volumineux. Y. Iliachenko a trouvé une démonstration du lemme précédent et du lemme de Bogdanov faisant appel non pas à des calculs mais à des considérations topologiques complexes (monodromie et formule de Picard-Lefchets) (voir Y. Iliachenko, Sur les zéros des intégrales abéliennes spéciales dans un domaine réel. Analyse fonctionnelle et ses applications, 11, 4 (1977), 78-79).

J. Cas d'une résonance d'ordre 4.

L'équation initiale s'écrit: $\dot{z} = \varepsilon z + Az\,|z|^2 + B\bar{z}^3$.

Supposons que $B \neq 0$. Alors par des dilatations et des rotations sur z et des dilatations du temps on ramène cette équation à la forme

$$\dot{z} = \varepsilon z + Az\,|z|^2 + \bar{z}^3.$$

On peut faire $\operatorname{Re} A \leqslant 0$ et $\operatorname{Im} A \leqslant 0$ en changeant le signe du temps et en remplaçant z par $\bar{z}$.

Voyons tout d'abord les points singuliers différents de zéro.

1°. *Bifurcations des points singuliers.* Pour étudier les bifurcations des points singuliers lorsque ε varie, on aura besoin de la construction auxiliaire suivante. Soit $z = re^{i\varphi}$ un point singulier. Alors

$$-\varepsilon/r^2 = A + N, \quad \text{où} \quad N = e^{-4i\varphi}.$$

Considérons donc le cercle de rayon 1 centré en A (fig. 146). La valeur de ε pour laquelle le point z est singulier est située sur l'arc opposé au rayon reliant le point 0 au point $A + N$ du cercle. En outre, plus le point du cercle est proche de 0, plus le module de ε est grand.

D'où il suit que les cas où $|A| < 1$ et $|A| > 1$ sont nettement distincts. Si $|A| < 1$, alors le zéro est contenu à l'intérieur du cercle. Dans ce cas quel que soit ε $(\neq 0)$, l'équation possède 4 points singuliers aux sommets d'un carré. Lorsque ε fait un tour autour du point 0, le carré des points singuliers pivote de 90° dans le sens inverse.

Si $|A| > 1$ il se forme sur le plan de la variable un angle limité par les prolongements des tangentes au cercle. Si ε est à l'intérieur de cet angle, on

a 8 points singluiers, s'il est à l'extérieur, aucun. Lorsque ε pivote d'un côté de l'angle à l'autre, 4 points singuliers apparaissent aux sommets d'un carré. Ce carré se partage immédiatement en deux. Ensuite les points singuliers voisins commencent à s'écarter l'un de l'autre. Lorsque ε se rapproche de l'autre côté de l'angle, chaque point singulier du premier carré meurt en heurtant le point du deuxième carré qui initialement était en retard de 90° sur lui. (De sorte qu'un carré de points singuliers pivote de 90° par rapport à l'autre.)

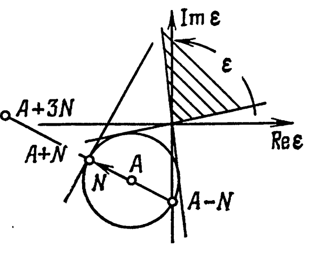


Fig. 146

2°. *Types de points singuliers de l'équation linéaire.* On commencera par un lemme qui nous permettra d'étudier facilement les types de points singuliers d'un champ de vecteurs sur le plan, dont l'écriture complexe est

$$\dot{\xi} = P\xi + Q\bar{\xi}.$$

L e m m e. *Le type du point singulier* 0 *ne dépend pas de l'argument de* Q. *Ce point est un col pour* $|P| < |Q|$, *un foyer pour* $|\operatorname{Im} P| > |Q|$ *et un nœud pour* $|\operatorname{Im} P| < |Q| < |P|$; *le foyer est stable pour* $\operatorname{Re} P < 0$ *et instable pour* $\operatorname{Re} P > 0$ (fig. 147).

◀ Lorsqu'on multiplie ξ par un nombre complexe λ, le coefficient P ne varie pas, quant au coefficient Q, il est multiplié par $\bar{\lambda}/\lambda$. On peut rendre Q

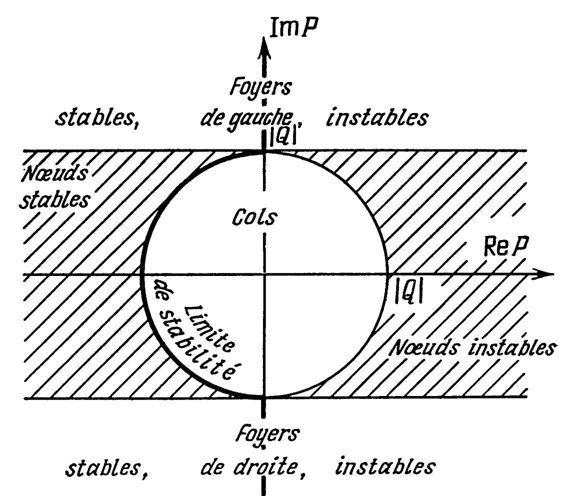


Fig. 147

quelconque par le choix de λ; ceci démontre la première proposition du lemme. Pour prouver la seconde, considérons le cas $Q = 1$. Supposons que $P = \alpha + i\beta$. Ecrivons la matrice de l'équation dans la base $(1, i)$:

$$M = \begin{pmatrix} \alpha+1 & -\beta \\ \beta & \alpha-1 \end{pmatrix}, \quad \operatorname{tr} M = 2\alpha, \quad \det M = \alpha^2 + \beta^2 - 1.$$

Donc les racines de l'équation caractéristique sont

$$\lambda_{1,2} = \alpha \pm \sqrt{1-\beta^2}.$$

Pour qu'elles soient réelles il faut que $|\beta| < 1$. Les racines sont de signe contraire pour $\alpha^2 + \beta^2 > 1$. Le lemme est démontré pour $Q = 1$. Les conditions d'existence d'un col, d'un nœud ou d'un foyer pour tout $|Q|$ découlent maintenant de la similitude: P et Q sont multipliées par un même nombre réel par les changements du temps.

La figure 147 qui indique nettement la disposition respective des nœuds, foyers et cols dans l'espace fonctionnel ne doit pas être perdue de vue lorsqu'on étudie les bifurcations des points singuliers sur le plan.

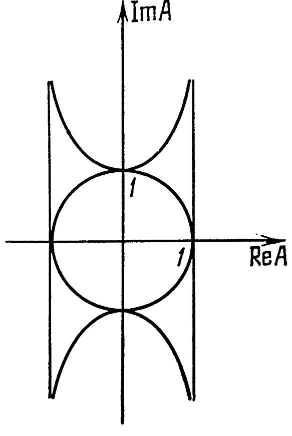


Fig. 148

3°. *Etude des cols.* Revenons à l'équation non linéaire initiale. Soit $z_0 = re^{i\varphi}$ un point singulier. Linéarisons l'équation en ce point. Supposons que $z = z_0 + \xi$. En conservant les termes du premier degré en ξ, $\bar{\xi}$ au second membre, on obtient

$$\dot{\xi} = P\xi + Q\bar{\xi}, \quad P = r^2(A - N), \quad |Q| = r^2(A + 3N).$$

L e m m e. *Si* $|A| < 1$, *alors tous les points singuliers sont des cols. Si* $|A| > 1$, *alors pour chaque* ε *le point singulier de plus petit module est un col, celui de plus grand module n'est pas un col.*

◀ D'après le lemme du 2° la condition d'existence d'un col s'écrit $|A - N| < |A + 3N|$. Considérons les points $A - N$ et $A + 3N$ (cf. fig. 146). Ces points sont symétriques par rapport à $A + N$, et, de plus, la droite qui les joint passe par le point A. La position du point 0 par rapport à la tangente menée au cercle par le point $A + N$ nous indique lequel de ces deux points est plus proche de 0. Si $|A| < 1$, alors 0 est toujours situé d'un même côté de la tangente (du même côté que $A - N$). Si $|A| > 1$, il faut déterminer sur lequel des deux arcs limités par les tangentes menées au cercle à partir du point 0 est situé $A + N$. L'arc le plus éloigné de zéro correspond aux cols et aux points singuliers proches de zéro (cf. 1°). ▶

4°. *Stabilité des points singuliers.* Les points singuliers correspondants à la partie du cercle tournée vers 0 peuvent être des nœuds et des foyers. La partie de l'arc qui est en contact avec la tangente menée à partir du point 0 correspond à un nœud, mais lorsque le point représentatif se déplace sur cet arc, ce nœud peut devenir un foyer et ce foyer peut changer de stabilité. Voyons à quelle condition a lieu ce changement de stabilité du foyer.

Le lemme du 2° et la formule du 3° nous disent que le changement de stabilité a lieu lorsque le point $A - N$ (qui est diamétralement opposé au point $A + N$ de notre cercle) coupe l'axe imaginaire, alors que le point $A + N$ est situé sur l'arc le plus proche de 0. La frontière séparant les points A pour lesquels ce phénomène a lieu est définie par la condition: le diamètre mené par le point d'intersection de notre cercle avec l'axe imaginaire est orthogonal à la tangente menée à partir du point 0. L'équation de la frontière s'écrit:

$$|\mathrm{Im}\, A| = (1 + \mathrm{Re}^2 A)/\sqrt{1 - \mathrm{Re}^2 A}.$$

La courbe correspondante sur le plan de la variable A coupe le cercle $|A| = 1$ aux points $A = \pm i$ et admet les droites $|\mathrm{Re}\, A| = 1$ pour asymptotes (fig. 148).

5°. *Comportement à l'infini.* On peut « négliger » le terme εz pour les grands q. En posant $w = z^2$, on obtient l'équation linéaire

$$\dot{w} = 2Aw + 2\bar{w}.$$

Appliquons le lemme du 2° à l'étude de cette équation. Pour $|A| < 1$ le point singulier sur le plan w est un col, pour $|A| > 1$ toutes les trajectoires sont attirées de l'infini vers un domaine fini si $\operatorname{Re} A < 0$.

6°. *Bifurcations du portrait de phases.* Si $|A| < 1$, la situation (fig. 149) est probablement la même que pour la résonance d'ordre 3 (cf. point G). Si $|\operatorname{Re} A| > 1$, il se passe probablement la même chose que pour les résonances

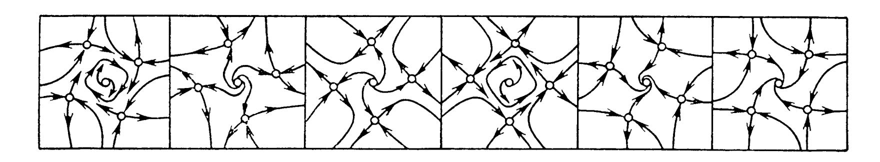


Fig. 149

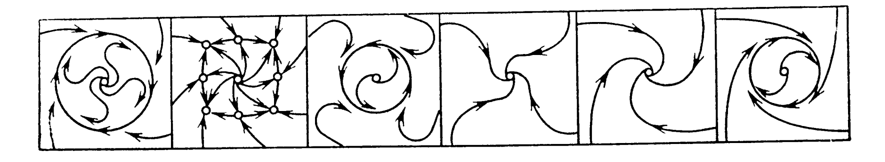


Fig. 150

d'ordre 5 et plus (fig. 150, 145) ; du reste dans ce cas les points singuliers peuvent naître en dehors du cycle. Comparer avec les exercices 1), 2), 3) de la page 318.

Les cas les plus difficiles sont $|\operatorname{Re} A| < 1$, $|A| > 1$.

7°. *Nouvelles normalisations et notations.* Pour étudier le cas $|\operatorname{Re} A| < 1$ il est utile de considérer l'asymptotique pour $|\operatorname{Im} A| \to \infty$. Au lieu de faire tendre $A \to \infty$ on peut faire tendre $B \to 0$ dans l'équation initiale. On adoptera les notations suivantes:

$$\varepsilon = \sigma + i\tau, \quad A = -i\alpha - \gamma, \quad B = \beta,$$

et l'on admettra que les grandeurs β, $\sigma = u\beta$, $\gamma = v\beta$ ($\beta \to 0$, $u \sim v \sim 1$) sont de petits paramètres du même ordre.

Choisissons les facteurs normalisant la dilatation des coordonnées, tels que $\alpha = 1$, $\tau = 1$.

Soient les coordonnées polaires symplectiques: $\rho = |z|^2/2$, $\varphi = \arg z$.

L'équation initiale prend la forme du système

$$\begin{cases} \dot\rho = 2\rho(\sigma - 2\gamma\rho + 2\beta\rho\cos 4\varphi), \\ \dot\varphi = \tau - 2\alpha\rho - 2\beta\rho\sin 4\varphi. \end{cases}$$

Soient encore le hamiltonien $H = \tau\rho - \alpha\rho^2 - \beta\rho^2 \sin 4\varphi$ et le potentiel $\Pi = \sigma\rho^2 - 4\gamma\rho^4/3$. Alors

$$\dot\rho = -H_\varphi + \Pi_\rho, \qquad \dot\varphi = H_\rho.$$

Dans le cas $\tau = \alpha = 1$, $\sigma = u\beta$, $\gamma = v\beta$ qui nous intéresse, on a

$$H = H_0 + \beta H_1, \ \Pi = \beta\Pi_1, \ H_0 = \rho - \rho^2.$$

Pour $\beta = 0$ on obtient un mouvement non perturbé (une rotation de fréquence H_ρ). Dans le domaine où $H_\rho \neq 0$ (c'est-à-dire lorsque ρ n'est pas voisin de 1/2), la principale perturbation est le fait du terme dissipatif $\beta\Pi_1$ et, pour $\rho \approx 1/2$, il faut tenir compte de βH_1.

8°. *Lemme sur l'effet d'une petite dissipation.* Considérons sur l'espace des phases l'équation $\dot{x} = v + \varepsilon w$, où v est un champ hamiltonien de hamiltonien H, w un champ potentiel, $w = \nabla\Pi$ (il s'agit d'une métrique définie par les coordonnées symplectiques p, q).

L e m m e. *La dérivée par rapport au petit paramètre ε de l'accroissement de H en un tour de trajectoire fermée $H = h$ est égale pour $\varepsilon = 0$ à*

$$\frac{d}{d\varepsilon}\,\delta H = \iint\limits_{G(h)} \Delta\Pi\, dp\, dq$$

où $G(h)$ est le domaine limité par la trajectoire.

◀ $$\delta H \approx \oint \dot{H}\, dt = \oint H_p(-H_q + \varepsilon\Pi_p) + H_q(H_p + \varepsilon\Pi_q)\, dt =$$
$$= \varepsilon \oint \Pi_p\, dq - \Pi_q\, dp = \varepsilon \iint \Delta\Pi\, dp\, dq.$$ ▶

En appliquant le lemme à notre équation on trouve

$$\delta H_0 = 2\rho(\sigma - 2\gamma\rho) = 2\beta\rho(u - 2v\rho).$$

Donc, en première approximation le cycle est donné par la formule

$$\rho = u/(2v) = \sigma/(2\gamma).$$

Cette méthode permet d'étudier ce système en dehors de la couronne dans laquelle ρ est proche de 1/2. Donc, le cas où σ est proche de γ doit être étudié séparément.

9°. *Etude du cas* $\rho \approx 1/2$. Faisons le changement $\rho = \frac{1}{2} + \sqrt{\beta}\, P$, $t\sqrt{\beta} = T$. Posons $\beta = 0$ après le changement. Pour équation approchée on obtient le système hamiltonien

$$dP/dT = w + \cos 4\varphi, \quad d\varphi/dT = -2P$$

de hamiltonien

$$H_{00} = P^2 + w\varphi + (\sin 4\varphi)/4$$

un pendule à moment de torsion). Ici $w = u - v = (\sigma - \gamma)/\beta$.

On a une fosse potentielle pour $|w| < 1$.

Dans les notations du 7° nous avons fait passer le champ $(\sigma - \gamma)\,\partial/\partial\rho$ de la partie potentielle du système dans la partie hamiltonienne. Donc, le nouveau hamiltonien et le nouveau potentiel s'écrivent

$$H = \rho - \rho^2 - \beta\rho^2 \sin 4\varphi - \beta w\varphi,$$
$$\Pi = \sigma\rho^2 - (4\gamma\rho^3)/3 - (\sigma - \gamma)\rho.$$

En appliquant le lemme du 8° on trouve

$$\delta H = \iint (2\sigma - 8\gamma\rho)\, dp\, d\varphi \text{ à l'intérieur de la fosse} = 2S(\sigma - 4\gamma\rho_0),$$

où S est la surface à l'intérieur de la fosse sur le plan (ρ, φ) et ρ_0 la coordonnée du centre de gravité de la fosse.

La condition de naissance d'un cycle est $\sigma = 4\gamma\rho_0$. Donc, il faut calculer ρ_0. Or $\rho_0 = \frac{1}{2} + \sqrt{\beta}\,\rho_1 + \ldots$. Donc, la condition de naissance d'un cycle devient $\sigma = 2\gamma + 4\gamma\sqrt{\beta}\rho_1 + \ldots$, $w = (\gamma/\beta) + 4\gamma\rho_1 + \ldots$.

Calculons ρ_1. L'équation exacte de la trajectoire fermée $H =$ const s'écrit dans les coordonnées P, φ

$$P^2 + \left(\frac{1}{2} + \sqrt{\beta}\, P\right)^2 \sin 4\varphi + w\varphi = h.$$

A toute valeur de φ dans la fosse correspondent deux valeurs de P telles que

$$P_1 + P_2 = -(\sqrt{\beta} \sin 4\varphi)/(1 - \beta \sin 4\varphi).$$

De cette formule il suit que le terme correctif désigné par ρ_1 est égal à

$$\rho_1 = -\frac{1}{2}\, \overline{\sin 4\varphi}$$

(la barre représente la « moyennisation » sur la trajectoire non perturbée pour $\beta = 0$).

La disposition de la fosse par rapport au maximum et au minimum de $\sin 4\varphi$ et sa variation, lorsque w varie (fig. 151), nous renseignent sur le comportement de ρ_1 qui nous permet de construire les modifications de la figure 152.

Le système de modifications représentées sur la figure 152 est réalisé pour des valeurs de $|\operatorname{Im} A|$ grandes en regard de $|B|$ pour $0 < |\operatorname{Re} A| < |B|$.

Il est entendu que les considérations précédentes ne peuvent se substituer aux démonstrations et ne constituent que l'ébauche de l'étude des bifurcations de la famille principale d'équation symétrique d'ordre 4.

Pour des cas de symétrie d'ordre $q \neq 4$ les résultats sont probablement connus depuis assez longtemps des spécialistes; en particulier, F. Takens les a annoncés en 1974 dans un travail polycopié non publié. Notre exposé est basé sur un article publié par l'auteur dans «Analyse fonctionnelle et ses applications», 1977, nº 2; des démonstrations détaillées ont été effectuées par F. Horozov.

K. Application de Poincaré.

Les applications de nos constructions à l'étude de la perte de la stabilité d'un cycle repose sur ce qui suit.

Lemme 1. *Considérons une application* $f\colon (\mathbb{R}^2, 0) \to (\mathbb{R}^2, 0)$ *possédant un point fixe* 0 *avec les valeurs propres* $e^{\pm 2\pi i p/q}$ *(et de bloc de Jordan d'ordre* 2 *si* $q = 1$ *ou* 2*). Alors* f^q *peut être représentée quel que soit* N*, dans un voisinage assez petit du point* 0*, par la somme* $f^q = g + h$*, où* $h(z) = O(|z|^N)$*, et* g *est une application du flot d'un champ de vecteurs invariant par un groupe cyclique fini de difféomorphismes* γ*, d'ordre* q*.*

On remarquera en particulier que g commute avec la rotation d'angle $2\pi/q$. L'application f^q ne commute généralement avec l'action d'aucun groupe fini et n'est pas incluse dans un flot. Le lemme 1 montre toutefois qu'au niveau des séries formelles f^q fait partie du flot et commute avec le groupe fini.

◀ La démonstration du lemme s'effectue suivant le schéma habituel de construction des formes normales de Poincaré-Dulac-Birkhoff (cf. chap. 5). ▶

Lemme 2. *Considérons une déformation* f_λ *d'une application* $f_0 = f$ *vérifiant les conditions du lemme* 1*. Alors quel que soit* N *l'itération* f_λ^q *peut être représentée dans un voisinage suffisamment petit du point* 0 *par la somme* $f_\lambda^q = g_\lambda + h_\lambda$*, où* $h_\lambda(z) = O(|z|^N)$*, et* g_λ *est*

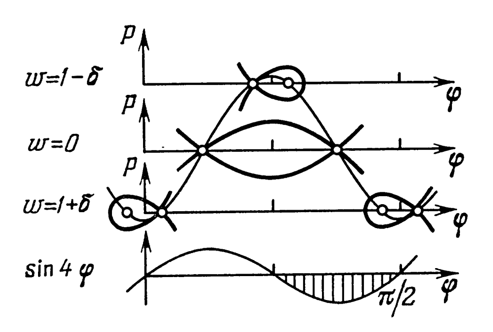


Fig. 151

Fig. 152

une transformation du flot d'un champ de vecteurs v_λ, *invariant par un groupe cyclique fini de difféomorphismes* γ_λ. *Ici* f_λ, g_λ, h_λ, v_λ *et* γ_λ *dépendent différentiablement du paramètre* λ *qui varie au voisinage de* 0.

◀ La démonstration est basée sur le fait que la réduction des termes de degrés $<N$ à la forme normale en négligeant les termes résonnants est réalisée par un difféomorphisme dépendant différentiablement d'un paramètre. ▶

En combinant le lemme 2 avec la description des bifurcations des flots des paragraphes précédents, on est renseigné sur la perte de stabilité du point fixe 0 de l'application f (ou d'un mouvement périodique pour lequel f est l'application de Poincaré).

R e m a r q u e. On peut également réduire directement à la forme normale une famille d'équations différentielles au voisinage d'une solution périodique (p, q)-résonnante dans l'espace d'un recouvrement à q feuillets. La méthode standard de Poincaré-Dulac-Birkhoff ramène alors la famille des champs de vecteurs 2π-périodiques en t à une somme d'un champ q-symétrique ne dépendant pas du temps et d'un résidu $O(|z|^N)$ de période $2\pi q$.

L. Discussion.

1°. Pour traduire les résultats obtenus dans le langage des bifurcations des solutions périodiques il faut remplacer les points fixes trouvés sur le plan par des trajectoires fermées dans l'espace, les séparatrices des points par les surfaces attractives et répulsives invariantes de ces trajectoires fermées, les cycles limites sur le plan par des tores invariants. La seule différence fondamentale résidera dans les modifications des séparatrices. Alors que sur le plan les séparatrices passent instantanément l'une à travers l'autre pendant les bifurcations, dans l'espace ce processus est durable et il se forme une figure homoclinique ou hétéroclinique *) (cf. fig. 101).

Dans l'espace, les tores invariants se détruisent avant que le cycle n'aborde le plan à une boucle de séparatrices: cependant, tous ces effets foncièrement tridimensionnels sont faibles (se déroulent à partir de termes de degré aussi grands que l'on veut dans les formes normales) par rapport aux effets bidimensionnels considérés plus haut.

2°. La considération de la perte de stabilité comme un phénomène à deux paramètres et non à un paramètre permet de comprendre facilement certaines circonstances semblant étonnantes.

Considérons une famille à deux paramètres dans laquelle pour paramètre on prend le multiplicateur lui-même. Traçons sur le plan des valeurs du paramètre le domaine d'existence des solutions périodi-

*) On appelle figure homoclinique (hétéroclinique) le réseau constitué sur un plan sécant par les traces concourantes des variétés invariantes attractive et répulsive d'une (de deux) trajectoire fermée.

ques qui se bouclent après avoir effectué q tours le long de la solution principale et p tours latéralement à cette solution. Ce domaine aborde le cercle unité au point $e^{2\pi i p/q}$ sous la forme d'une cuspide (dont la largeur à une distance σ du cercle est de l'ordre de $\sigma^{(q-2)/2}$, pour $q > 4$, voir point F). Donc la courbe générique coupe une infinité de cuspides (fig. 153) au voisinage du cercle unité dans le plan du multiplicateur.

Donc, dans une famille générique à un paramètre dans laquelle le cycle perd sa stabilité sans résonance forte, à un instant proche de l'instant de perte de stabilité naissent et meurent une infinité de cycles de période longue.

La démonstration de ce fait qui ne dépend pas de l'existence d'une résonance faible au moment de la perte de la stabilité a été

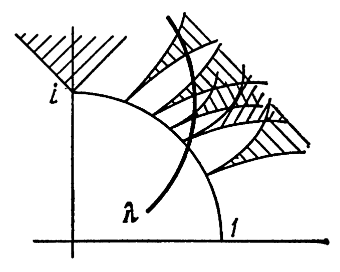


Fig. 153

effectuée par V. Koziakine (V. Koziakine, Subfurcation des oscillations périodiques, DAN SSSR, 232, 1 (1977), 25-27 (en russe)).

3°. Lorsqu'on étudie des courbes génériques sur les diagrammes des bifurcations de fortes résonances, cités au point F, on peut décrire des suites de modifications qui sont universelles mais qui paraissent non locales du point de vue « à un paramètre ».

Par exemple, pour $q = 2$ les événements peuvent se dérouler comme suit. Le cycle stable perd doucement sa stabilité avec formation d'un tore qui commence aussitôt à se déformer de telle sorte que la forme de son méridien se rapproche d'un huit; à l'approche du centre du huit (où se trouve un cycle instable) l'ensemble attractif, tout en restant proche du tore dont le méridien a pratiquement pris la forme d'un huit, se détruit au voisinage de la séparatrice homoclinique (Y. Neimark).

Dans ce cas, la trajectoire effectue un tour tantôt autour d'une moitié, tantôt autour de l'autre moitié du tore détruit en sautant de l'une à l'autre d'une manière qui semble aléatoire.

Cette description ressemble au phénomène observé dont une expérience numerique de Guertsenstein et Schmidt (S. Guertsenstein, V. Schmidt, Développement non linéaire et interaction des perturbations d'amplitude finie pendant l'instabilité convective d'une couche plane en rotation, DAN SSSR, 225, 1 (1975), 59-62 (en russe)).

§ 36. Modifications de la topologie en cas de résonances

Les résonances entre les valeurs propres de la partie linéaire d'un champ de vecteurs en un point stationnaire empêchent de choisir les coordonnées linéarisant ce champ. Même s'il n'y a pas de résonance, mais les valeurs propres sont proches des valeurs résonnantes, les séries de Poincaré peuvent être divergentes et le système ne peut alors être linéarisé par un changement analytique de coordonnées.

Ceci étant le type topologique du portrait du système n'est pas généralement modifié dans un voisinage réel du point stationnaire en cas de résonance. Par exemple, si les parties réelles de toutes les valeurs propres sont négatives, le point stationnaire est attractif et indépendamment des résonances le système est topologiquement équivalent au système linéaire standard.

En cas de résonance la topologie est modifiée, mais généralement dans le plan complexe.

Le système générique n'est pas résonnant. Les résonances sont inévitables dans une famille à un paramètre. Donc en étudiant l'influence des résonances sur la modification de la topologie on doit considérer des familles à un paramètre de champs de vecteurs. D'après ce qui précède, on admettra que les coordonnées dans l'espace des phases, le temps et le paramètre sont des nombres complexes.

Résonances dans le domaine de Poincaré.

Considérons des trajectoires complexes dans le voisinage du point singulier O. Ces trajectoires forment un feuilletage de feuilles à deux dimensions réelles avec une singularité en 0. Pour comprendre la structure de cette singularité, coupons ce feuilletage par une sphère de petit rayon centrée en l'origine des coordonnées.

Supposons que la partie linéaire du système considéré est diagonale dans les coordonnées $(z_1, \ldots, z_n)$: $\dot{z}_j = \lambda_j z_j + \ldots$, $j = 1, \ldots, n$.

T h é o r è m e. *Si une collection de valeurs propres* $\{\lambda_j\}$ *appartient au domaine de Poincaré, alors chaque sphère* $|z_1|^2 + \ldots + |z_n|^2 = r^2$ *de rayon assez petit coupe transversalement le feuilletage.*

◀ Considérons d'abord un système linéaire. On a

$$dr^2 = \sum \bar{z}_j\, dz_j + z_j\, d\bar{z}_j = A\, dt + \bar{A}\, \overline{dt}, \quad A = \sum |z_j|^2 \lambda_j.$$

La condition de transversalité à la sphère consiste en ce que la 1-forme dr^2 ne doit pas s'annuler sur un plan tangent à une feuille. Or la forme $A\, dt + \bar{A}\, \overline{dt}$ est nulle uniquement pour $A = 0$. La relation $A = 0$ n'est réalisée dans le cas de Poincaré (et dans ce cas

seulement) pour aucun $z \neq 0$. Donc, le théorème est démontré dans le cas linéaire: les feuilles intersectent la sphère sous un angle $\alpha(z)$ non nul.

Etudions le minimum α_0 de l'angle $\alpha(z)$ sur la sphère $|z| = r$. La grandeur α_0 ne dépend pas de r (puisque $\alpha(cz) = \alpha(z)$). Donc, $\alpha(z) \geqslant \alpha_0 > 0$ pour tous les $z \neq 0$.

Revenons maintenant au système non linéaire. L'angle entre les champs de directions du système non linéaire et sa partie linéaire est petit avec $|z|$. Donc dans un voisinage assez petit de 0 il est inférieur à α_0 et les trajectoires du système non linéaire intersectent la sphère transversalement. ▶

Corollaire. *Les intersections des trajectoires complexes et d'une sphère de rayon assez petit forment un feuilletage à une dimension sans point singulier sur cette sphère. Les feuilletages obtenus sur toutes les sphères de rayons assez petits sont difféomorphes. Le type différentiable du feuilletage sur la sphère ne change pas par les déformations de la sphère tant qu'elle est transversale aux trajectoires complexes.*

Donc, le feuilletage à deux dimensions étudié au voisinage d'un point singulier est homéomorphe à un cône au-dessus d'un feuilletage à une dimension sur la sphère. Ce feuilletage sur la sphère est la décomposition en trajectoires du champ de vecteurs sur la sphère (puisque la sphère et le feuilletage complexe sont orientables).

Remarque. En cas de non-résonance le théorème de Poincaré nous dit que dans des coordonnées convenables le système est linéaire dans un voisinage assez petit du point singulier. D'où il suit que le type différentiable du feuilletage sur la sphère dans le cas non résonnant est le même que pour le système linéaire.

On conclut donc que le type différentiable d'un feuilletage défini par un système non linéaire sur la sphère est le même que pour la partie linéaire du système non seulement dans ce voisinage de l'origine des coordonnées, où convergent les séries de Poincaré, mais loin de ce voisinage.

En effet, à l'approche de la résonance, le rayon du domaine de convergence des séries de Poincaré tend vers zéro alors que celui du domaine de transversalité reste minoré. Donc, la variation du feuilletage à une dimension sur une sphère de petit rayon fixe (ne dépendant pas du paramètre) nous renseigne sur le passage par la résonance dans le système complexe.

B. Résonance $\lambda_1 = 2\lambda_2$.

A titre d'exemple, voyons comment varie la topologie du feuilletage sur S^3 au passage par la résonance $\lambda_1 = 2\lambda_2$ dans le système

$$\dot{z}_1 = \lambda_1 z_1 + \ldots, \quad \dot{z}_2 = \lambda_2 z_2 + \ldots$$

On se trouve dans le domaine de Poincaré si le rapport $\lambda = \lambda_1/\lambda_2$ n'est pas un nombre réel négatif. Considérons tout d'abord le feuilletage sur S^3 correspondant à la partie linéaire du système.

Les séparatrices $z_1 = 0$, $z_2 = 0$ intersectent la sphère suivant de grands cercles qui sont des cycles du système sur S^3. Leur coefficient d'enlacement est égal à 1.

Si λ n'est pas un nombre réel (cas d'un foyer), toutes les autres trajectoires du feuilletage sur la sphère se déroulent d'un cycle et s'enroulent sur un autre. Etudions les applications de Poincaré des cycles.

On remarquera que ces applications peuvent être supposées holomorphes. En effet, elles sont équivalentes au sens des difféomorphismes réels aux applications complexes de Poincaré envoyant une transversale holomorphe à la séparatrice dans elle-même. Donc, elles deviennent holomorphes par un choix convenable d'une structure complexe sur une transversale à deux dimensions réelles à un cycle de S^3. D'où il suit aussi que les multiplicateurs de nos cycles sont égaux respectivement à $e^{\pm 2\pi i\lambda}$ et $e^{\pm 2\pi i\lambda^{-1}}$.

Les feuilletages sur S^3 correspondant à tous les foyers sont homéomorphes mais non difféomorphes: $\lambda^2 + \lambda^{-2}$ est un invariant du difféomorphisme.

Si λ est un nombre positif réel (cas d'un nœud), on se trouve également dans le domaine de Poincaré. Dans ce cas la partie de S^3, comprise entre deux cycles enlacés, est décomposée en tores à deux dimensions portant des hélices partout denses, de nombre de rotation identique sur tous ces tores.

Etudions maintenant le système non linéaire. Dans le cas d'un foyer la résonance est impossible, donc le feuilletage sur une sphère est difféomorphe au feuilletage décrit plus haut pour le système linéaire. Il en va de même pour un nœud non résonnant, c'est-à-dire pour tous les $\lambda > 0$, à l'exception du cas où λ ou $1/\lambda$ est un nombre entier.

Voyons, par exemple, la résonance $\lambda = 2$. La forme normale de Poincaré s'écrit alors:

$$\dot{z}_1 = \lambda_1 z_1 + c z_2^2, \qquad \dot{z}_2 = \lambda_2 z_2.$$

Ce système ne possède pour $c \neq 0$ qu'une séparatrice, et le feuilletage sur S^3 qu'un seul cycle. Remplaçons les λ proches de 2 par une valeur non réelle. Le système obtenu sur S^3, d'un côté, est proche d'un système résonnant à un cycle, de l'autre, est difféomorphe au système étudié précédemment et construit d'après un foyer linéaire et, par suite, possède deux cycles de coefficient d'enlacement égal à 1. On démontre que l'un de ces cycles, disons C_1 pour fixer les idées, est proche de l'unique cycle C du système résonnant. L'autre cycle C_2 est situé sur un tore fin de ligne axiale C_1 et se boucle après avoir effectué deux tours le long de C_1 et un tour sur le méridien (de sorte que les coefficients d'enlacement de C_2 et C_1 sont égaux à 1). Donc, la modification du système sur S^3 au voisinage de la résonance $\lambda = 2$ consiste en la séparation d'une trajectoire périodique double d'une trajectoire périodique de valeurs propres (—1, —1).

Remarque. L'exposé précédent fait suite à une note de l'auteur « Remarques sur les singularités de codimension finie dans les systèmes dynamiques complexes », Analyse fonctionnelle et ses applications 3, 1 (1969), 1-6. Les résultats de cette note ont été par la suite généralisées par J. Guckenheimer, N. Kuiper, N. Ladis, Y. Iliachenko et autres. Les résultats relatifs au type topologique du feuilletage défini au voisinage d'un point singulier dans l'espace complexe d'un système linéaire sont les plus complets.

Etudions notamment le cas où l'espace des phases est à trois dimensions et le triangle des valeurs propres contient le zéro strictement à l'intérieur. Il s'avère que le type topologique du feuilletage est déterminé dans l'espace complexe par un triplet de valeurs inverses des valeurs propres, considérées comme un triplet de vecteurs sur le plan réel (c'est-à-dire considérées aux applications linéaires réelles près de l'espace du plan réélifié d'une variable complexe). Dans le cas multidimensionnel le type réel de la collection des inverses des valeurs propres définit de façon biunivoque le type topologique du feuilletage complexe au voisinage d'un point singulier du système linéaire si le zéro appartient à l'enveloppe convexe des valeurs propres et si leurs quotients deux à deux

sont non réels (cf. C. Camacho, N. Kuiper, J. Palis, C. R. Acad. Sci. Paris, 282 (1976), 959-961; N. Ladis, Invariants topologiques des flots linéaires complexes, Equations différentielles, 12, 12 (1976), 2159-2169; Y. Iliachenko, Remarques sur la topologie des points singuliers des équations différentielles dans le domaine complexe et théorème de Ladis, Analyse fonctionnelle et ses applications, 11,2 (1977), 28-38).

C. Déformations verselles dans le cas de Poincaré.

Considérons un champ de vecteurs (différentiable) analytique de point singulier O. Supposons que ce point singulier est du type de Poincaré, c'est-à-dire l'enveloppe convexe de la collection de valeurs propres ne contient pas le point 0.

Théorème. *Le germe d'un champ de vecteurs analytique (resp. holomorphe, différentiable) en un point singulier de type de Poincaré possède une déformation verselle analytique (resp. holomorphe, différentiable) à nombre fini de paramètres, constituée de champs de vecteurs polynômiaux.*

En d'autres termes:

Une famille locale de champs de vecteurs analytique (resp. holomorphes, différentiables) de point singulier O (de type de Poincaré est analytiquement (resp. holomorphement, différentiablement) équivalent au voisinage du point O à une famille constituée de polynômes assez longs de Taylor de ces champs au point O.

◀ Le point singulier est non dégénéré par hypothèse, donc, il dépend différentiablement des paramètres. Par suite, on peut le ramener en l'origine des coordonnées par un changement de variables différentiable et dépendant différentiablement des paramètres. Supposons que les valeurs propres sont simples. Alors on peut choisir la base propre dépendante différentiablement des paramètres. Dans le système de coordonnées obtenu, la famille d'équations différentielles correspondant à notre famille de champs s'écrit

$$\dot{x}_k = \lambda_k(\varepsilon)\, x_k + \ldots, \quad k = 1, \ldots, n.$$

Utilisons la méthode de Poincaré (chap. 5) pour éliminer seulement les termes qui sont non résonnants pour $\varepsilon = 0$. Les changements de variables dépendent alors différentiablement du paramètre.

Comme les valeurs propres appartiennent au domaine de Poincaré, les résonances sont en nombre fini et la convergence de cette procédure se démontre sans peine.

On obtient un système de coordonnées dans lequel les seconds membres de toutes les équations de la famille sont des polynômes. ▶

Le cas où les seconds membres sont différentiables un nombre fini (ou infini) de fois se traite également sans peine. Pour les détails on pourra se référer à l'article de N. Brouchlinskaïa (N. Brouchlinskaïa), Théorème de finitude pour des familles de champs de vecteurs au voisinage d'un point singulier de type de Poincaré, Analyse fonctionnelle, 5, 3 (1971), 10-15) dans lequel est également étudié le cas des valeurs propres multiples.

D. Matérialisation des résonances.

La réduction à la forme normale dans le domaine de Siegel soulève des difficultés liées aux petits diviseurs. Ceci n'empêche pas toutefois que la structure topologique soit parfois simple. Par exemple, un col ordinaire présente la même structure topologique pour un quotient rationnel et un quotient irrationnel des valeurs propres. Le même phénomène a lieu dans le domaine de Poincaré. Les résonances peuvent n'exercer aucune influence sur la topologie du portrait de phase.

On peut se demander naturellement pourquoi les résonances qui ne se manifestent pas d'un point de vue topologique empêchent la réduction analytique (ou assez lisse) à la forme normale. Il est utile de ne pas perdre de vue le comportement des résonances en théorie des perturbations des mouvements quasi périodiques.

Considérons l'équation différentielle sur un tore T^n à n dimensions

$$\dot{\theta} = \omega + \varepsilon \ldots, \ \bmod 2\pi \in T^n, \qquad \omega \in \mathbb{R}^n, \ \varepsilon \ll 1. \qquad (*)$$

La résonance $(\omega, k) = 0$ modifie (pour le moins en l'absence de perturbation, c'est-à-dire pour $\varepsilon = 0$) les propriétés topologiques du système: les trajectoires sont partout denses dans un tore non pas à n dimensions comme dans le cas non résonnant, mais à $(n - 1)$ dimensions. Par exemple, pour $n = 2$, en cas de résonance sont engendrés généralement des régimes périodiques structurellement stables (des cycles limites stables et instables sur le tore). Il est clair que l'existence de tels cycles entrave la réduction des équations à la forme normale $\dot{\theta} = \omega$, ordinaire pour le cas non résonnant.

La démonstration de l'inexistence des intégrales premières du problème des trois corps, donnée par Poincaré, repose sur des raisonnements proches.

On peut supposer que l'influence des résonances sur la divergence dans le problème local étudié plus haut revêt le même caractère, mais est liée à des modifications de la topologie du feuilletage formé par les trajectoires non pas sur un domaine réel mais sur un domaine complexe. Même si elle n'exerce absolument aucune influence sur la partie réelle de l'espace des phases, une telle modification empêche à priori une réduction analytique et une réduction C^r-différentiable.

A noter que le système $\dot{x}_k = \lambda_k x_k + \ldots$ se ramène à la forme $(*)$ par la substitution $x = e^{i\theta}$ (les ω réels correspondent à des λ imaginaires purs). Les méthodes ordinaires de recherche de cycles limites du système $(*)$ nous amènent à considérer la première intégrale $\rho = e^{i(\theta, k)}$ du système non perturbé; dans des notations du système initial figure la grandeur $\rho = x^k$. L'équation de première approximation pour la variété invariante correspondant à la résonance se déduit à partir de la relation

$$\dot{\rho} = \rho\,[(k, \lambda) + (k, c)\,\rho + \ldots].$$

On obtient formellement

$$\rho \approx -\frac{(k,\ \lambda(\varepsilon))}{(k,\ c(\varepsilon))}.$$

Cependant nos séries sont généralement divergentes et notre conclusion doit être justifiée.

Pour $n = 2$ on peut asseoir nos raisonnements par une démonstration rigoureuse de l'existence d'un cycle limite complexe qui au voisinage de la résonance possède l'asymptotique indiquée (A. Piartly, Naissance de variétés invariantes complexes au voisinage d'un point singulier d'un champ de vecteurs dépendant des paramètres, Analyse fonctionnelle, 6, 4 (1972), 95-96).

Au moment de la résonance, lorsque $(k, \lambda) = 0$, le cycle (une trajectoire non simplement connexe complexe) se rapproche des séparatrices complexes du point singulier. Le chemin homotope à zéro de ce cycle disparaît en cas de résonance et fusionne avec la position d'équilibre. Un cas particulier est la naissance (ou la mort) d'un cycle à partir de la position d'équilibre pendant la perte de stabilité (voir § 33). Dans ce cas, $k = (1, 1, 0, \ldots)$, $\lambda_1 + \lambda_2 = 0$ et le phénomène est entièrement observable dans un domaine réel (cf. fig. 127). Dans les autres cas (même pour la même résonance, par exemple, dans le cas d'un col) la topologie des trajectoires réelles est susceptible de ne pas être modifiée par la résonance.

La différence entre la topologie des trajectoires complexes de l'équation (ou de la famille) et celle des trajectoires de sa forme normale est un obstacle à la réduction analytique à la forme normale. Bien plus, si cette différence (ce qui est généralement le cas) est définie d'après un jet d'ordre fini, elle fait obstacle non seulement à une réduction analytique à la forme normale, mais aussi à une réduction C^r-différentiable. Par exemple, la divergence des séries réductrices dans le cas, où le quotient des valeurs propres est trop bien approché par des nombres rationnels, peut être expliquée par l'existence dans tout voisinage d'un point stationnaire de cycles limites complexes provenant de résonances voisines d'ordre supérieur: la forme normale du système ne possède pas de tels cycles, donc la réduction à une forme normale est nécessairement divergente.

L'étude de la divergence des séries de Poincaré est loin de son terme. Les démonstrations de la divergence, antérieures au travail de Piartly (Poincaré, Siegel, Bruno) sont basées sur le calcul de l'accroissement des coefficients et ne déterminent pas les causes de la divergence de même que le calcul des coefficients de la série de arc tg z établit la divergence pour $|z| > 1$, mais ne détermine pas les causes: les singularités pour $z = \pm i$.

A. Piartly a établi les résultats suivants:

1°. Le passage par la résonance $k_1\lambda_1 + k_2\lambda_2 = 0$ dans $\mathbb{C}^2$ donne lieu dans le cas général à l'isolation des séparatrices du point singulier d'une variété invariante dont l'équation en première approxima-

tion est de la forme $z_1^{k_1}z_2^{k_2} = \varepsilon$, où ε caractérise l'écart par rapport à la résonance, et z_1, z_2 sont les coordonnées dans l'espace des phases.

2°. Un résultat analogue a été établi pour la même résonance dans $\mathbb{C}^n$ sous des conditions restrictives sur les autres valeurs propres.

3°. Pour λ_1 et λ_2 « anormalement commensurables », le cas générique est l'existence d'un nombre infini de variétés invariantes correspondant à des résonances différentes dans un voisinage quelconque du point singulier, ce qui entraîne la divergence des séries de Poincaré.

Le travail de A. Piartly s'appuie sur la méthode de E. Hopf. A. Bruno a proposé d'autres démonstrations et généralisations des deux premiers résultats de A. Piartly (A. Bruno, Formes normales des équations différentielles à petits paramètres, Mathemat. Zametki, **16**, 3 (1974), 407-414; Variétés invariantes analytiques, DAN SSSR, **216**, 2 (1974), 253-256; Ensembles analytiques intégraux DAN SSSR, **220**, 6 (1975), 1255-1258).

E. Résonance entre trois valeurs propres.

La résonance suivante par ordre de difficulté

$$k_1\lambda_1 + k_2\lambda_2 + k_3\lambda_3 = 0,$$

où le triangle λ_1, λ_2, λ_3 contient le point 0 a également été étudiée par A. Piartly et A. Bruno. On a également prouvé ici qu'une variété invariante se séparait des séparatrices du point singulier pour $n = 3$. Supposons que (z_1, z_2, z_3) sont les coordonnées dans l'espace des phases, ε le paramètre de la déformation (la résonance correspond à $\varepsilon = 0$). Alors dans l'espace muni des coordonnées (z, ε), les variétés invariantes constituées de trajectoires partout denses dans ces variétés forment une hypersurface holomorphe dont l'équation en première approximation est de la forme $z_1^{k_1}z_2^{k_2}z_3^{k_3} = \varepsilon$ dans un système de coordonnées convenable.

A. Piartly a démontré que pour les $(\lambda_1, \lambda_2, \lambda_3)$ « anormalement commensurables », qui constituent le triangle de Siegel, il existe dans le cas général dans un voisinage quelconque de la position d'équilibre $0 \in \mathbb{C}^3$ une infinité de variétés invariantes de la forme décrite, correspondant à diverses résonances.

Dans la suite, on verra que la présence dans un voisinage du point non résonnant $0 \in \mathbb{C}^3$ d'une portion assez grande de variété invariante résonnante fait obstacle à la convergence des séries de Poincaré-Siegel dans ce voisinage. Donc du résultat de Piartly il suit que pour $(\lambda_1, \lambda_2, \lambda_3)$ « anormalement commensurables » le cas générique est la divergence des séries indiquées dans un voisinage quelconque de l'origine des coordonnées.

F. Cas où le temps est discret.

Considérons un autodifféomorphisme local $A: \mathbb{C}^n \to \mathbb{C}^n$ au voisinage du point fixe 0. Désignons par $(\lambda_1, \ldots, \lambda_n)$ les valeurs propres du linéarisé de A en 0.

La théorie précédente se généralise à ce cas avec les modifications suivantes.

Résonances: $\lambda_k = \lambda_1^{m_1} \ldots \lambda_n^{m_n}$ $(m_i \geqslant 0, \ \sum m_i \geqslant 2)$.

Domaine de Poincaré: les $|\lambda_s|$ sont tous >1 ou tous <1.

Domaine de Siegel: il existe $|\lambda_i| \geqslant 1$ et $|\lambda_j| \leqslant 1$.

Remarque. A un champ de vecteurs linéaire de $\mathbb{C}^n$ correspond une application linéaire de $\mathbb{C}^{n-1}$ (« l'application de Poincaré »). Plus exactement, supposons que le champ définit une équation différentielle

$$\dot{z}_1 = \alpha_1 z_1, \ \ldots, \ \dot{z}_n = \alpha_n z_n,$$

où $\alpha_n \neq 0$. Considérons les solutions vérifiant pour $t = 0$ des conditions initiales sur le plan $z_n = 1$. La valeur prise par la solution pour $t = 2\pi i/\alpha_n$ appartient au même plan de $\mathbb{C}^{n-1}$. Les valeurs propres de l'application $A: \mathbb{C}^{n-1} \to \mathbb{C}^{n-1}$ sont $\lambda_s = e^{2\pi i \alpha_s/\alpha_n}$ $(s = 1, \ldots, n-1)$.

On peut construire d'une façon analogue l'application de Poincaré (sous des hypothèses subsidiaires faibles) dans le cas non linéaire. Donc, les résultats obtenus sur les variétés invariantes et les bifurcations pour les applications se généralisent aux champs des vecteurs.

Du reste, dans la plupart des cas il vaut mieux utiliser ce lien entre les équations et les applications comme procédé heuristique pour supposer les résultats dans un domaine d'après les résultats obtenus dans un autre, quant aux démonstrations, il est commode de les effectuer indépendamment l'une de l'autre dans les deux situations.

G. Bifurcation des variétés invariantes d'un difféomorphisme. La résonance entre les trois valeurs propres du champ de vecteurs

$$k_1\alpha_1 + k_2\alpha_2 + k_3\alpha_3 = 0$$

correspond dans le cas discret à une résonance de la forme

$$\lambda_1^{m_1}\lambda_2^{m_2} = 1,$$

et, de plus, dans le domaine de Siegel $m_1 > 0$, $m_2 > 0$, $|\lambda_1| \neq 1 \neq |\lambda_2|$. On suppose que $|\lambda_1| > 1$, et $|\lambda_2| < 1$.

Les résultats de Piartly et de Bruno traduisent dans ce cas l'existence de variétés invariantes constituées de trajectoires partout denses formant une surface de l'espace muni des coordonnées (ε, z_1, z_2), dont l'équation commence à partir des termes de la forme $\varepsilon = z_1^{m_1} z_2^{m_2}$. Ici ε représente l'écart par rapport à la résistance, (z_1, z_2), des coordonnées convenables (dépendant différentiablement de ε) dans l'espace des phases. Pour $\varepsilon \neq 0$ la variété invariante construite est homéomorphe à un cylindre. Les coefficients d'enlacement du cercle directeur de ce cylindre et des axes de coordonnées $\mathbb{C}^2$ sont égaux à m_1 et m_2.

On montrera que l'existence d'une portion assez grande d'une telle variété invariante résonnante dans un voisinage d'un point fixe non résonnant du difféomorphisme fait obstacle à la linéarisation de ce difféomorphisme dans ce voisinage. Donc, lorsque n'importe quel voisinage du point fixe comprend des variétés résonnantes, les séries linéarisantes sont partout divergentes.

Topologiquement, toutes les applications de $\mathbb{C}^2$, telles que $|\lambda_1| > 1 > |\lambda_2|$ sont équivalentes l'une à l'autre. En particulier, elles sont linéarisables et possèdent beaucoup de cylindres invariants. Cependant les cylindres invariants holomorphes sont, comme nous le verrons, très rares.

H. Translations locales.

On se propose de lier à une variété invariante résonnante de l'application $\mathbb{C}^2 \to \mathbb{C}^2$ une courbe elliptique plongée dans une surface holomorphe. Cette surface est la variété des orbites de notre application (ou la variété des trajectoires de l'équation différentielle initiale dans $\mathbb{C}^3$). Pour définir exactement la variété des orbites on aura besoin de la terminologie suivante.

Soit le cylindre $S^1 \times \mathbb{R}$. On appelle *translation standard* du cylindre l'adjonction de 1 à la deuxième coordonnée. Supposons que D_0 est un domaine du cylindre contenant $S^1 \times [0, 1]$. La restriction de la translation standard à D_0 définit un difféomorphisme $t: D_0 \to D_1 = t(D_0)$. On remarquera que l'intersection $D_0 \cap D_1$ contient le cercle $S^1 \times 1$. Soit D la réunion $D_0 \cup D_1$.

Supposons que M est une variété à deux dimensions, M_0 et M_1 des domaines de cette variété, $f: M_0 \to M_1$, un homéomorphisme.

D é f i n i t i o n. On dit que l'homéomorphisme f est une *translation locale* si existe un homéomorphisme $h: M \to D$ envoyant M_0 dans D_0, M_1 dans D_1 et f dans t.

Soient M une courbe complexe, $f: M_0 \to M_1$ une translation locale holomorphe. Alors le collage de chaque point $z \in M_0$ avec son image $f(z) \in M_1$ définit une courbe complexe compacte homéomorphe à un tore, c'est-à-dire une courbe elliptique $\Gamma = M/f$.

◀ La démonstration est évidente. ▶

Considérons maintenant le produit direct du cylindre par le plan $\Pi = (S^1 \times \mathbb{R}) \times \mathbb{R}^2$. Soient $T: \Pi \to \Pi$ une translation de 1 le long de $\mathbb{R}$, E_0 un voisinage de $D_1 \times 0$, $E_1 = TE_0$, $E = E_0 \cup E_1$.

Supposons que N est une variété réelle différentiable à quatre dimensions, $M \subset N$ une sous-variété à deux dimensions, N_1 et N_2 des domaines de $N = N_1 \cup N_2$, $F: N_1 \to N_2$ un homéomorphisme.

D é f i n i t i o n. On dit que l'homéomorphisme f est une *translation locale de N le long de M* si existe un homéomorphisme H: $N \to E$ envoyant N_0 dans E_0, N_1 dans E_1, F dans T et M dans D.

Supposons que N est une surface complexe, $M \subset N$ une courbe complexe, $F: N_0 \to N_1$ une translation locale holomorphe le long de M.

Alors le collage de chaque point $z \in N_0$ et de son image $F(z) \in N_1$ définit une surface complexe holomorphe $\Sigma = N/F$, qui est un voisinage de la courbe elliptique $\Gamma = M/F$.

◀ La démonstration est évidente. ▶

I. Construction d'une courbe elliptique d'après une variété invariante résonnante d'une application linéaire.

Soit $A : \mathbb{C}^2 \to \mathbb{C}^2$ une application linéaire de valeurs propres λ_1, λ_2 $|\lambda_1| > 1 > |\lambda_2|$. Supposons que les valeurs propres vérifient la relation de résonance $\lambda_1^{m_1}\lambda_2^{m_2} = 1$, où m_1 et m_2 sont premiers entre eux. Alors le cylindre d'équation $z_1^{m_1}z_2^{m_2} = 1$ (où z_1 et z_2 sont des coordonnées dans la base propre) est invariant par A.

La restriction de A à ce cylindre définit une translation holomorphe. En effet, uniformisons la courbe $\lambda_1^{m_1}\lambda_2^{m_2} = 1$ avec le paramètre $\Lambda \neq 0$ par la formule $\lambda_1 = \Lambda^{m_2}$, $\lambda_2 = \Lambda^{-m_1}$; introduisons sur le cylindre le paramètre $Z \neq 0$ tel que $z_1 = Z^{m_2}$, $z_2 = Z^{-m_1}$. Alors A agit sur le cylindre comme $Z \mapsto \Lambda Z$. Cette transformation est une translation holomorphe puisque $|\Lambda| > 1$. La courbe elliptique correspondante est $\mathbb{C}^*/\{\Lambda\} \cong \mathbb{C}/(2\pi\mathbb{Z} + \omega\mathbb{Z})$, où $\Lambda = e^{i\omega}$.

On remarquera que dans le cas d'une résonance cette application linéaire possède toute une famille à un paramètre de cylindres invariants holomorphes $z_1^{m_1}z_2^{m_2} = c$, $c \neq 0$. Les courbes elliptiques construites d'après ces cylindres sont isomorphes.

Le quotient d'un voisinage convenable d'un tel cylindre (ou d'une partie assez grande d'un tel cylindre) par l'action de A est un voisinage de la courbe elliptique sur une surface complexe. Cette surface est le produit direct de la courbe elliptique par $\mathbb{C}$. En effet, les homothéties $z \mapsto kz$ définissent la projection sur la courbe elliptique, l'application $z \mapsto z_1^{m_1}z_2^{m_2}$ la projection sur le deuxième facteur.

En particulier, l'indice de self-intersection de la courbe elliptique est nul dans la surface construite.

J. Construction d'une courbe elliptique d'après une variété invariante résonnante d'une application non linéaire.

Soit $A(\varepsilon) : U \to \mathbb{C}^2$ une application biholomorphe de $U \subset \mathbb{C}^2$ dans $\mathbb{C}^2$, dépendant holomorphement du paramètre ε. On suppose que ε varie dans un voisinage de 0 dans $\mathbb{C}$ et que l'origine des coordonnées de $\mathbb{C}^2$ est invariante par toutes les applications $A(\varepsilon)$.

Soient λ_1 et λ_2 les valeurs propres de la linéarisée de l'application $A(0)$ en 0. On supposera que $|\lambda_1| > 1 > |\lambda_2|$ et que $\lambda_1^{m_1} = \lambda_2^{-m_2}$, où m_1 et m_2 sont premiers entre eux.

Pour la famille générique A, au passage par la résonance ($\varepsilon = 0$) un cylindre résonnant holomorphe invariant se sépare des séparatrices du point fixe (voir point G). Fixons un ε assez petit et considérons la restriction de $A(\varepsilon)$ à ce cylindre. On vérifie que $A(\varepsilon)$ induit sur la partie correspondante du cylindre une translation holomorphe locale. [Ceci résulte du fait que:

1°. En première approximation le cylindre a pour équation $z_1^{m_1} z_2^{m_2} = c\ (\varepsilon)$;

2°. $A\ (\varepsilon)$ est proche de la linéarisée de $A\ (0)$ en 0.

3°. La linéarisée de $A\ (0)$ dans 0 agit sur le cylindre comme une translation locale (cf. point I).]

Donc, pour $|\ \varepsilon\ |$ petit l'application $A\ (\varepsilon)$ définit une courbe elliptique $\Gamma\ (\varepsilon)$ plongée dans une surface $\Sigma\ (\varepsilon)$. Le groupe des homologies de la courbe $\Gamma\ (\varepsilon)$ contient le cercle distingué (l'image du cercle directeur du cylindre). La courbe $\Gamma\ (\varepsilon)$ peut être représentée par

$$\Gamma\ (\varepsilon) \approx \mathbb{C}/(2\pi\mathbb{Z} + \omega\ (\varepsilon)\ \mathbb{Z}),$$

où 2π correspond au cercle distingué. La fonction $\omega\ (\varepsilon)$ tend vers ω_0 pour $\varepsilon \to 0$. Les formules du point I entraînent que $\lambda_1 = e^{i\omega_0 m_2}$, $\lambda_2 = e^{-i\omega_0 m_1}$.

L'indice de self-intersection de la courbe Γ dans la surface Σ est nul. Donc, topologiquement Σ est le produit direct de Γ par un disque. Cependant analytiquement Σ n'est pas nécessairement un produit direct. Bien plus,

1) Σ peut ne pas être un fibré holomorphe au-dessus de Γ. Le voisinage de Γ dans Σ peut ne pas admettre d'applications holomorphes sur Γ identiques sur Γ. Ce sera le cas, par exemple, lorsqu'un voisinage de Γ dans Σ contiendra une famille de courbes elliptiques avec diverses valeurs du module ω.

2) Γ peut ne pas admettre dans Σ de déformations distinctes de translations de Γ le long de Γ. Ceci aura lieu, par exemple, si le fibré normal de Γ dans Σ est analytiquement non trivial.

Au paragraphe 27 on trouvera des résultats sur la classification de Σ.

K. Non-linéarisation d'une application dans un domaine contenant un cylindre résonnant.

Théorème. *Si une courbe elliptique est indéformable dans son voisinage, alors l'application A n'est pas linéarisable par un changement de variables biholomorphe au voisinage du point* 0 *contenant la partie du cylindre invariant holomorphe, nécessaire à la construction de la courbe* Γ.

◀ En effet, on peut toujours déformer un cylindre invariant holomorphe d'une application linéaire par une homothétie voisine de l'identité. ▶

Donc, nous avons obtenu une majoration du rayon de convergence des séries de Poincaré-Siegel aux moyens d'un cylindre invariant holomorphe de fibré normal non trivial de la courbe elliptique correspondante.

Théorème (Y. Iliachenko). *Si une application linéaire possède un cylindre invariant holomorphe dont le cercle*

directeur possède avec les axes propres les coefficients d'enlacement (m_1, m_2), *et la courbe elliptique correspondante est* $\mathbb{C}/(2\pi\mathbb{Z} + \omega\mathbb{Z})$ (*en outre*, 2π *correspond à la directrice du cylindre*), *alors les valeurs propres sont égales à* $\lambda_1 = e^{i\omega m_2}$, $\lambda_2 = e^{-i\omega m_1}$ *et, par suite, on a la résonance* $\lambda_1^{m_1}\lambda_2^{m_2} = 1$.

◀ Soient (z_1, z_2) les coordonnées propres. Les formes différentielles dz_k/z_k sur le cylindre sont holomorphes et invariantes par l'application, donc, elles définissent des formes holomorphes sur la courbe elliptique.

Calculons les intégrales de ces formes le long des génératrices du groupe d'homologies du tore. L'une de ces génératrices correspond au cercle directeur γ sur le cylindre. Pour elle on a

$$\oint dz_1/z_1 = 2\pi i m_2, \qquad \oint dz_2/z_2 = 2\pi i m_1,$$

puisque le coefficient d'enlacement de γ avec l'axe $z_1 = 0$ est égal à m_2, et avec l'axe $z_2 = 0$, égal à m_1. La deuxième génératrice correspond au segment δ, qui relie le point z à son image Az dans la surface du cylindre. Pour elle on a $\int dz_1/z_1 = \ln \lambda_1$, $\int dz_2/z_2 = \ln \lambda_2$ (ces relations définissent les branches des logarithmes). Or toutes les formes holomorphes sur la courbe elliptique sont proportionnelles. Donc,

$$\omega/2\pi = \ln \lambda_1/2\pi i m_2 = -\ln \lambda_2/2\pi i m_1. \quad ▶$$

C o r o l l a i r e. *Soient A un difféomorphisme local de point fixe* 0, $\lambda_{1,2}$ *ses valeurs propres. Supposons qu'un voisinage U du point fixe contient un cyclindre holomorphe, que A est une translation holomorphe sur ce cylindre et* ω *la période de la courbe elliptique correspondante. Si* λ *et* ω *ne sont pas reliés par la relation* $\lambda_1 = e^{i\omega m_2}$, $\lambda_2 = e^{-i\omega m_1}$ (*où* m_1 *et* m_2 *sont les coefficients d'enlacement avec les séparatrices du point singulier*), *alors le difféomorphisme A n'est pas équivalent analytiquement à un difféomorphisme linéaire dans le domaine U.*

L. Divergence des séries de Poincaré.

Les résultats précédents entraînent le

T h é o r è m e. *Si un voisinage aussi petit que l'on veut d'un point fixe non résonnant* 0 *d'un difféomorphisme local de* $\mathbb{C}^2 \to \mathbb{C}^2$ *contient des cylindres holomorphes sur lesquels le difféomorphisme agit comme une translation locale, alors ce difféomorphisme n'est analytiquement équivalent à un difféomorphisme linéaire en aucun voisinage du point fixe* 0 (*et, par suite, les séries de Poincaré sont partout divergentes*).

A. Piartly a établi qu'une telle accumulation des cylindres invariants au point singulier est un cas générique pour les applica-

tions dont les valeurs propres sont « anormalement bien approchées par des valeurs résonnantes ». Donc pour de telles valeurs propres la divergence des séries de Poincaré est un cas générique en tout voisinage du point fixe.

Les résultats de ce paragraphe se généralisent aisément aux champs de vecteurs dans $\mathbb{C}^3$ au voisinage d'un point singulier du type de Siegel. La matérialisation de la résonance $m_1\lambda_1 + m_2\lambda_2 + m_3\lambda_3 = 0$ est une courbe elliptique. Les points de cette courbe sont les trajectoires du champ situé sur la surface résonnante invariante $z_1^{m_1}z_2^{m_2}z_3^{m_3} = c\varepsilon + \ldots$

M. Bifurcations de courbes elliptiques sur des surfaces complexes. La théorie des bifurcations de courbes elliptiques à indice de self-intersection nul sur des surfaces complexes est analogue à celle des bifurcations des variétés invariantes d'équations différentielles.

Une courbe elliptique et son fibré normal sont donnés sur une surface par un couple de nombres complexes (ω, λ) (cf. § 27) : on les obtient à partir de l'axe des complexes φ et du plan des variables complexes (r, φ) par les collages

$$(r, \varphi) \sim (r, \varphi + 2\pi) \sim (\lambda r, \varphi + \omega).$$

Un fibré est dit *résonnant* s'il devient analytiquement trivial par passage à un revêtement cyclique de base à nombre fini de feuillets.

Les fibrés résonnants correspondent dans l'espace des couples (λ, ω) aux hypersurfaces $\lambda^n = e^{ikw}$. Il s'avère que lorsque le couple (composé de la courbe elliptique et de la surface) varie continûment une courbe elliptique revêtant topologiquement la courbe considérée s'approche d'elle au moment du passage par la résonance. Donc, la résonance est matérialisée par la bifurcation d'une courbe elliptique multiple.

Soit une famille à un paramètre de couples $\Gamma(\varepsilon) \subset \Sigma(\varepsilon)$. Supposons que $\varepsilon = 0$ correspond à la résonance $\lambda^n = e^{ik\omega}$. Il se trouve que l'équation de la courbe approchante est de la forme $r^n e^{ik\varphi} = \varepsilon$ (par un choix d'un paramètre ε convenable, et un changement de coordonnées (r, φ) convenable, dépendant de ε; on admet que la résonance correspond à $\varepsilon = 0$ et qu'il n'existe pas de résonances d'ordre inférieur: $\lambda^m e^{il\varphi} \neq 1$ pour $0 < m < n$).

Voyons comment on déduit l'équation de la courbe approchante à l'aide des séries formelles. En reprenant les raisonnements du § 27 on ramène le collage à la forme

$$\begin{cases} r \\ \varphi \end{cases} \mapsto \begin{cases} r\lambda(1 + \alpha\varepsilon + aw + A), \\ \varphi + \omega + \beta\varepsilon + bw + B, \end{cases}$$

où α, β, a, b sont des constantes, $w = r^n e^{ik\varphi}$; A et B sont des séries sur les puissances de ε et w commençant par les termes d'ordre 2. Cette substitution transforme w en $w(1 + \gamma\varepsilon + cw + C)$, où $\gamma = n\alpha + ik\beta$, $c = na + ikb$, C sont des termes d'ordre 2 sur ε et w.

L'équation $\gamma\varepsilon + cw + C = 0$ définit la courbe approchante. Pour la famille générique on a $\gamma \neq 0$, $c \neq 0$. Un changement convenable des coordonnées ε et r ramène cette équation à la forme $r^n e^{ik\varphi} = \varepsilon$.

On étudie la convergence comme dans les travaux de Piartly et Bruno cités.

R e m a r q u e. On vérifie sans peine que la condition $\lambda^n e^{ik\varphi} = 1$ signifie très exactement que le fibré normal est analytiquement trivial au-dessus d'un recouvrement cyclique à nombre fini de feuillets d'une courbe elliptique.

Du point de vue topologique, tous les fibrés considérés sont triviaux. En particulier, la courbe elliptique approchante pendant la résonance se projette

(non holomorphement) sur une courbe Γ (ε) « le long de la direction ». Cette projection est un revêtement cyclique topologique à nombre fini de feuillets du tore. C'est ce revêtement sur lequel le fibré normal devient trivial au moment du passage par la résonance.

Dans le cas où $(n, k) = d > 1$ (mais où il n'existe pas de résonance d'ordre inférieur, $\lambda^m e^{il\omega} \neq 1$ pour $0 < m < n$) la courbe approchante n'est pas connexe. Elle est alors constituée de composantes dont chacune est un revêtement topologique à (n/d)-feuillets du tore initial.

N. Divergence des séries linéarisantes.

Les séries ramenant le collage à la forme normale divergent pour certains fibrés non résonnants (c'est-à-dire pour certains couples λ, ω).

La bifurcation des courbes matérialisant la résonance permet d'« expliquer » la divergence des séries linéarisant le collage de la manière suivante. Supposons qu'un couple (λ, ω) n'est pas résonnant, mais est très proche d'une résonance. Alors dans un petit voisinage de la courbe elliptique initiale, il existe en principe une autre courbe elliptique, à savoir la courbe matérialisant la résonance. Si le couple (λ, ω) est assez proche d'un nombre infini de résonances, dans un voisinage arbitrairement petit de la courbe elliptique initiale, il existe une infinité de courbes matérialisant les diverses résonances et recouvrant cycliquement la courbe initiale.

D'autre part, le fibré normal de la courbe initiale est non résonnant. Les fibrés normaux non résonnants de degré zéro *) ne possèdent de sections au-dessus d'aucun recouvrement cyclique à un nombre fini de feuillets d'une courbe elliptique. Donc, le fibré normal de la courbe elliptique initiale ne contient pas de courbes elliptiques revêtant cycliquement la courbe initiale. Par suite, aucun voisinage de la courbe initiale sur la surface n'est appliqué biholomorphement sur un voisinage de la section nulle du fibré normal; donc les séries sont divergentes pour les collages génériques si le couple (λ, ω) est trop bien approché par des couples résonnants.

R e m a r q u e. Il existe une analogie entre les sous-variétés complexes compactes de variétés analytiques et les cycles limites des équations différentielles: de même qu'un cycle limite est susceptible de disparaître par une petite déformation du champ seulement si l'opérateur de monodromie possède une valeur propre égale à 1, de même une courbe elliptique sur une surface possédant un indice de self-intersection nul ne disparaît pas par de petites déformations de la surface ambiante si le fibré normal est analytiquement non trivial. La formulation générale suivante est due à F. Bogomolov: une sous-variété compacte d'une variété complexe ne disparaît pas par une petite déformation de la variété ambiante si les cohomologies de dimension 1 du faisceau normal sont triviales.

§ 37. Classification des points singuliers

Dans ce paragraphe nous renoncerons au point de vue « universel » pour considérer non pas des familles, mais des systèmes d'équations différentielles au voisinage d'un point singulier d'un champ de vecteurs en admettant qu'il existe des dégénérescences de codimension aussi grande que l'on veut. D'un point de vue général l'étude de singularités aussi compliquées présente un intérêt très limité, puisque ces dégénérescences possèdent une codimension élevée et se rencontrent rarement.

*) On appelle degré d'un fibré l'indice de self-intersection de la section nulle de ce fibré.

Cependant la connaissance des traits généraux de singularités arbitraires est intéressant même dans les cas compliqués qui sont hors de portée des techniques actuelles.

En particulier, la connaissance de la nature des pathologies que l'on peut rencontrer dans les cas de codimension élevée est utile ne serait-ce que pour ne pas se dépenser inutilement à la recherche de choses inexistantes, comme ce fut le cas pour la recherche, par exemple, de critères algébriques de stabilité au sens de Liapounov ou de stabilité asymptotique, ainsi que des critères algébriques dans le problème du centre-foyer pour des racines nulles de l'équation caractéristique.

Pour montrer de quoi il retourne nous allons nous pencher sur un exemple très simple qui se traite entièrement.

A. Points singuliers de fonctions sur une droite.

Soit f une fonction réelle différentiable au voisinage du point $x = 0 \in \mathbb{R}$. Si le point 0 n'est pas critique, la fonction est différentiablement équivalente dans son voisinage à une fonction linéaire ($f(x) = x + c$).

On sait bien également ce qui se passe dans le cas critique: si $f'(0) = 0$ le comportement de la fonction est défini par le signe de $f''(0)$, etc.

Pour fixer les idées, considérons le problème des conditions de minimum de la fonction au point 0. On peut donner la réponse sous la forme suivante: l'espace J^k des k-jets des fonctions en 0 se décompose en trois parties:

$$J^k = \mathrm{I} \cup \mathrm{II} \cup \mathrm{III};$$

I sont les jets garantissant un minimum,
II les jets garantissant une absence de minimum,
III les jets qui ne permettent pas de conclure à un minimum.

Les jets de type I et II sont dits *suffisants*, ceux du type III, *réceptifs*.

Les ensembles I, II, III jouissent des deux propriétés suivantes.

1°. *Semi-algébricité*. Chacun des ensembles I, II, III est une sous-variété semi-algébrique de l'espace des jets J^k.

Un *ensemble semi-algébrique* dans $\mathbb{R}^N$ est défini comme la réunion finie de sous-ensembles définis chacun par un système fini d'équations et d'inéquations polynômiales.

Si les inéquations sont superflues, l'ensemble est dit algébrique. Une propriété utile des ensembles semi-algébriques est traduite par le théorème suivant (pour la démonstration voir A. Seidenberg, A new decision method for elementary algebra, Ann. of Math., Ser. 2, **60** (1954), 356-374; E. Gorine, Sur les propriétés asymptotiques des polynômes et des fonctions algébriques, UMN, **16**, 1 (1961), 91-118).

Principe de Tarski-Seidenberg. *L'image d'un ensemble semi-algébrique par une application polynômiale est semi-algébrique.*

Voici une formulation équivalente mais plus faible:

La projection d'un ensemble semi-algébrique dans un sous-espace est un ensemble semi-algébrique.

A noter que la projection d'un ensemble algébrique peut déjà ne pas être algébrique mais être seulement ensemble semi-algébrique (exemple: la projection d'une sphère sur un plan).

2°. *Définition de presque finitude.* Pour $k \to \infty$ la codimension de l'ensemble des jets réceptifs III $\subset J^k$ tend vers l'infini.

En d'autres termes, les jets réceptifs dans J^k sont définis par un nombre de conditions croissant avec k. Il s'avère en fin de compte que l'ensemble des fonctions pour lesquelles le problème, qui consiste à déterminer si 0 est un minimum local, n'est soluble pour aucun nombre de termes de la série de Taylor, est maigre: il est de codimension infinie dans l'espace fonctionnel.

B. Autres exemples.

Le problème analogue pour des fonctions de plusieurs variables ne possède pas un algorithme simple: si la seconde différentielle s'annule il faut faire appel aux suivantes et l'on est conduit à des problèmes du type de la classification des courbes algébriques, des surfaces, etc. Cependant même ici les partitions J^k = I $\cup$ II $\cup$ III de l'espace des k-jets des fonctions dans $\mathbb{R}^n$ sont semi-algébriques et presque finies bien qu'il soit pratiquement impossible d'écrire sous une forme explicite les équations et les inéquations en les coefficients de Taylor pour quelques n et k grands. L'existence de ces équations et inéquations peut être établie à partir du théorème de Tarski-Seidenberg dont la démonstration renferme également un algorithme pour leur écriture (ce théorème généralise la théorie de Sturm).

L'intervalle initial de la classification a été calculé explicitement et s'est trouvé mystérieusement lié à la classification des polyèdres réguliers, avec les groupes de Coxeter, Weyl et Lie des séries A_k, D_k, E_k, avec les fonctions automorphes, les triangles sur le plan de Lobatchevski, avec les singularités des caustiques, des fronts d'ondes et des intégrales oscillantes de la méthode de phase stationnaire (cf. aperçu: V. Arnold, Points singuliers de fonctions différentiables et leurs formes normales; UMN, **30**, 5 (1975), 3-65 et la bibliographie de cette aperçu; V. Vassiliev, Asymptotique des intégrales exponentielles, diagramme de Newton et classification des points de minimum, Analyse fonctionnelle et ses applications, **11**, 3 (1977), 1-11.

L'exemple suivant est un problème de classification topologique des germes d'applications différentiables. En 1964, R. Thom a annoncé le théorème de semi-algébricité et de presque finitude dans ce problème; la démonstration a été donnée par A. Vartchenko

(A. V a r t c h e n k o, Propriétés topologiques locales des applications différentiables, Izvestia AN SSSR, Sér. Matém. 38, 5 (1974), 1037-1090).

C. Points singuliers de champs de vecteurs.

Revenons au problème de la classification topologique des points singuliers, des champs de vecteurs. Au départ la situation semble aussi simple que pour les fonctions. Les points singuliers non dégénérés sont classés d'après le nombre de valeurs propres dans le demi-plan de gauche. L'espace des 1-jets est décomposé en un nombre fini de domaines en fonction du nombre de racines dans le demi-plan de gauche. Chacun de ces domaines est un ensemble semi-algébrique dans l'espace des jets; les inéquations polynômiales qui le définissent peuvent être explicitées (conditions de Routh-Hurwitz, cf., par exemple, F. G a n t m a c h e r, Théorie des matrices, M., « Naouka », 1967).

Les 1-jets réceptifs forment une sous-variété semi-algébrique de codimension un, séparant les domaines correspondant à un nombre différent de racines dans le demi-plan de gauche. Dans les paragraphes précédents nous avons vu de nombreux exemples traitant des cas dégénérés au passage aux 2-jets, etc. Donc, on a l'impression qu'ici on peut aller aussi loin que l'on veut et que seule la complexité des calculs et la pléthore de variantes ne permettent pas de procéder à une classification algébrique dans les cas d'une codimension arbitrairement grande. Cependant la situation est tout autre (V. A r n o l d, Insolubilité algébrique du problème de stabilité au sens de Liapounov et du problème de la classification topologique des points singuliers d'un système analytique d'équations différentielles, Analyse fonctionnelle, 4, 3 (1970), 1-9).

La propriété de semi-algébricité n'est même pas réalisée dans un problème aussi simple que le problème de différentiation d'un centre et d'un foyer pour des racines nulles de l'équation caractéristique (A. Bruno, Y. Iliachenko; voir Y. I l i a c h e n k o, Insolubilité algébrique et solubilité presque algébrique du problème du centre-foyer, Analyse fonctionnelle, 6, 3 (1972), 30-37). Donc, il n'existe pas d'algorithme algébrique *) pour les problèmes de stabilité et de classification topologique.

Il reste encore à espérer qu'il existe un algorithme non algébrique, c'est-à-dire que la propriété de presque finitude ait tout de même lieu: l'ensemble des germes dont le type topologique (ou la stabilité) n'est défini d'après aucun polynôme de Taylor est peut être de codimension infinie. La vérification de ce fait est assez compliqué; il importe d'indiquer exactement ce qu'on entend par codimension:

*) Tout récemment L. Hazine et E. Cnol ont établi l'insolubilité algébrique du problème de stabilité dans le cas de deux couples de valeurs propres imaginaires pures de résonance 3 : 1. Ce cas correspond à une sous-variété de codimension 3 dans l'espace fonctionnel.

les ensembles de l'espace des k-jets dont il faut déterminer la codimension ne sont pas algébriques et l'on peut éprouver des difficultés de nature ensembliste. R. Thom présume que la réponse à cette question est négative.

D. Structure des ensembles de réceptivité.

Le problème de presque finitude se rattache au comportement des ensembles de jets réceptifs dans l'espace de k-jets J^k pour $k \to \infty$. La structure des ensembles de réceptivité pour k fixe semble plus accessible à l'étude. Fixons un $(k - 1)$-jet réceptif d'un champ de vecteurs en 0 et considérons l'espace J de tous les k-jets de $(k - 1)$-jet donné. Pour fixer les idées, on étudiera le problème de la stabilité asymptotique. Alors l'espace J est partagé en sous-ensembles (éventuellement vides): le sous-ensemble I des k-jets suffisants pour la stabilité, le sous-ensemble II des k-jets suffisants pour l'instabilité et un reste III de jets réceptifs (pour la classification topologique le nombre de sous-ensembles est plus élevé). Une position raisonnable du problème du critère de stabilité consiste à établir les propriétés des sous-ensembles I, II et de la frontière qui les sépare. Par exemple, la transcendance de la frontière traduit l'inexistence d'un critère algébrique de stabilité. On se demande quel est le degré de complexité de la structure de cette frontière. Peut-elle, par exemple, posséder un nombre infini de composantes de connexité? (*Idem* pour les ouverts I, II). Les points des sous-ensembles I et II peuvent-ils être mélangés comme les points rationnels et irrationnels?

On ne connaît pas d'exemples de cette nature, cependant il est à craindre que la situation ne se présente justement ainsi dans les cas singuliers de codimension assez grande dans un espace multidimensionnel.

Il existe un lien étroit entre le problème local du comportement des trajectoires au voisinage d'un point singulier de $\mathbb{R}^n$ et le problème global des équations différentielles définies par un système polynômial dans l'espace projectif $\mathbb{R}P^{n-1}$. Dans le travail consacré à l'insolubilité algébrique cité plus haut, ce lien a été utilisé pour établir la transcendance de la frontière de stabilité dans l'espace des jets du problème local à partir de la transcendance de la surface de naissance des cycles limites dans l'espace des coefficients du système polynômial sur le plan projectif. Or dans le cas global multidimensionnel il peut apparaître des phénomènes bien plus complexes que les cycles limites, par exemple, des systèmes sur un tore avec des nombres de rotation commensurables et incommensurables mélangés ou des domaines dans l'espace fonctionnel, libre de systèmes structurellement stables. Tous ces phénomènes sont réalisés dans les systèmes polynômiaux dans l'espace projectif et chacun d'eux est susceptible de contribuer à l'embrouillement de la frontière de stabilité dans l'espace J.

EXEMPLES DE PROBLÈMES D'EXAMEN

Les épreuves par écrit qui durent 4 heures comportent 15 questions rattachées entre elles. La note attribuée à chaque question est indiquée entre crochets. Elle est communiquée aux candidats au début de l'épreuve.

Variante 1.

$$\ddot{x} = -\sin x + \varepsilon \cos t. \qquad (1)$$

I. Soit $\varepsilon = 0$.

1) Linéariser l'équation donnée au point $x = \pi$, $\dot{x} = 0$ [1].

2) Etudier la stabilité de cette position d'équilibre [1].

3) Trouver la matrice de Jacobi de la transformation du flot dans le temps $t = 2\pi$ au point $x = \pi$, $\dot{x} = 0$ [3].

4) Trouver la dérivée de la solution vérifiant la condition initiale $x = \pi$, $\dot{x} = 0$ par rapport aux paramètres ε pour $\varepsilon = 0$ [5].

5) Tracer les graphes de la solution vérifiant la condition initiale $x = 0$, $\dot{x} = 2$ et de sa dérivée par rapport à t [3].

6) Trouver cette solution [3].

II. Supposons que (2) est l'équation aux variations le long de la solution indiquée au point 5.

7) L'équation (2) admet-elle des solutions non bornées [8] ?

8) L équation (2) possède-t-elle des solutions non nulles bornées pour $-\infty < t < \infty$ [8] ?

9) Trouver le wronskien du système fondamental des solutions de l'équation (2), sachant que $W(0) = 1$ [5].

10) Expliciter l'équation (2) et l'intégrer [10].

11) Trouver les valeurs et les vecteurs propres de l'opérateur de monodromie pour l'équation aux variations le long de la solution vérifiant la condition initiale $x = \pi/2$, $\dot{x} = 0$ [16].

12) Prouver que l'équation (1) possède une solution 2π-périodique, dépendant différentiablement de ε et égale à $x = \pi$ pour $\varepsilon = 0$ [6].

13) Calculer la dérivée de cette solution par rapport à ε pour $\varepsilon = 0$ [6].

III. Soit l'équation $u_t + uu_x = -\sin x$.

14) Trouver l'équation des caractéristiques [2].

15) Trouver la plus grande valeur t pour laquelle la solution du problème de Cauchy avec $u|_{t=0} = 0$ est prolongeable sur $[0, t[$ [8].

Variante 2.

I. Supposons qu'un champ de vecteurs dans un espace à trois dimensions possède un point singulier nul et en outre une valeur propre du point singulier est nulle et les deux autres, imaginaires pures.

1) Réduire à la forme normale les termes d'ordre 1 du développement des composantes du champ en série de Taylor au point 0 [1].

2) *Idem* pour les termes d'ordre 2 [3].
3) *Idem* pour les termes d'ordre quelconque [8].
4) Moyenniser le système sur la rotation rapide définie par la partie linéaire du champ [12].

II. Soit donnée une famille de champs dépendant différentiablement d'un paramètre et contenant le champ du point I pour la valeur nulle du paramètre.

5) Réduire à sa forme la plus simple le polynôme de Taylor des champs de la famille en 0 par un difféomorphisme dépendant différentiablement des paramètres variables au voisinage de 0 [10].

6) Appliquer à la même famille la moyennisation sur la rotation rapide définie par la partie linéaire du champ initial [20].

III. Dans l'espace des 1-jets des champs de vecteurs dans $\mathbb{R}^3$ considérons la variété des jets au points singuliers possédant une valeur propre nulle et deux imaginaires pures.

7) Trouver la codimension de cette variété [2].

8) Trouver la condition de transversalité de la famille écrite sous la forme trouvée dans l'exercice 5) à la variété indiquée [8].

9) Etudier les bifurcations des points singuliers dans des familles génériques à deux paramètres, transversales à la variété indiquée [10].

10) Etudier les bifurcations des cycles nées de ces points singuliers [15].

11) Etudier l'existence et la différentiabillité de la trajectoire reliant ces points singuliers [15].

IV. Repérons une droite passant par le point 0 dans le plan. Un difféomorphisme du plan est distingué s'il conserve la droite distinguée. Un champ de vecteurs est distingué s'il est tangent à la droite distinguée en tous ses points. Soit donné un tel champ ayant en l'origine des coordonnées un point singulier. Supposons que les deux valeurs propres du linéarisé du champ en 0 sont égales à zéro.

12) Réduire le polynôme de Taylor du champ au point 0 à sa forme la plus simple par un difféomorphisme acceptable [12].

13) Réduire une famille de champs acceptables qui est une déformation du champ donné, à la forme normale formelle par des difféomorphismes formels acceptables dépendant formellement de paramètres variables au voisinage du zéro [16].

14) Etudier les bifurcations des points singuliers dans les familles génériques obtenues à partir des formes normales de l'exercice 13) en négligeant les termes d'ordre supérieur [18].

15) Appliquer les résultats des problèmes 12), 13), 14) à l'étude des bifurcations des portraits de phase des champs dont une valeur propre est nulle, et deux, imaginaires pures [25].

Exercices supplémentaires

1) Soit $\dot{z} = t + Az \mid z \mid^2 + \bar{z}^3$. Pour $\mid \operatorname{Re} A \mid > 1$ le nombre de cycles limites est $\leqslant 1$.

Indication. Diviser le champ par $z\bar{z}$; div $P(z, \bar{z}) = 2\operatorname{Re}(\partial P/\partial z)$.

2) Soit $A = (3 + i)/\sqrt{2}$. Si $\arg \varepsilon = 5\pi/4$, alors la séparatrice singulière de chaque col-nœud est confondue avec la séparatrice non singulière du suivant.

Indication. Lorsque le col se confond avec un nœud, l'équation se ramène à la forme:

$$\dot{w} = e^{i\theta} [Rw(|w|^2 - 1) + i(\bar{w}^2 - w^2)], \quad A = (R - i) e^{i\theta}.$$

Si $R = 2$ et $\theta = \pi/4$, alors les séparatrices sont des droites.

3) Etudier les courbes séparant les domaines de valeurs de A, où les points singuliers coïncident sur, à l'intérieur et à l'extérieur du cycle lorsque $\arg \varepsilon$ varie.

Indication. Une variation de θ fait pivoter le champ. Les courbes sont disposées approximativement comme les 4 paraboles: $a^2 = 2(\pm b \pm 1)$, $A = a + ib$.

4) Pour $|\operatorname{Re} A|$ petits et $1 < |\operatorname{Im} A| < c \approx 4{,}11$ l'équation 1) possède pour ε convenable deux cycles limites, à l'intérieur desquels sont situés 9 points singuliers et pour $|\operatorname{Im} A| > c$, un seul point (A. Neishtadt). La frontière rappelle une ellipse d'axes de longueur 3, 11 et 2.

5) Dans le système $\dot{x} = x(\alpha + ax + by)$, $\dot{y} = y(\beta + cx + dy)$ il n'y a pas de cycles limites.

Indication. Lorsque le foyer perd sa stabilité le système possède une intégrale première: le produit des puissances de trois fonctions linéaires (N. Baoutine).

INDEX ALPHABÉTIQUE